PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/29, 15/82, C12Q 1/68

A2

(11) International Publication Number: WO 97/06259

(43) International Publication Date: 20 February 1997 (20.02.97)

(21) International Application Number: PCT/EP96/03480

(22) International Filing Date: 6 August 1996 (06.08.96)

(30) Priority Data:
95401849.5 7 August 1995 (07.08.95) EP
(34) Countries for which the regional or international application was filed: GB et al.

(71) Applicant *(for all designated States except US)*: KEYGENE N.V. [NL/NL]; Agrobusiness Park 90, P.O. Box 216, NL-6700 AE Wageningen (NL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ZABEAU, Marc [BE/BE]; Erpelsteeg 70, B-9000 Gent (BE). VOS, Pieter [NL/NL]; Dorpsstraat 22, NL-3927 BD Renswoude (NL). SIMONS, Guus [NL/NL]; Zoetendaal 11, NL-6715 JL Ede (NL).

(74) Agents: GUTMANN, Ernest et al.; Ernest Gutmann-Yves Plasseraud S.A., 3, rue Chauveau-Lagarde, F-75008 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: RESISTANCE AGAINST WILT INDUCING FUNGI

(57) Abstract

The invention relates to a nucleic acid comprising the I-2 resistance gene which, when present and expressed in a plant, is capable of conferring said plant resistance against wilt inducing fungi. The DNA sequence is at least part of the DNA sequence provided in the figure or any DNA sequence homologous thereto. The invention also relates to a gene product encoded by the I-2 gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

RESISTANCE AGAINST WILT INDUCING FUNGI

FIELD OF THE INVENTION

The present invention relates to resistance genes, DNA constructs, micro-organisms, plant cells and plants comprising said resistance genes. Furthermore the invention relates to genetically transformed plants which are resistant against wilt inducing fungi. In addition, the invention relates to probes, and primers for the identification of the resistance genes and diagnostic kits comprising said probes and/or primers. Finally, the invention relates to polypeptides encoded by said resistance genes and the use of said polypeptides.

BACKGROUND OF THE INVENTION

Plant pathogens such as fungi are responsible for substantially losses of plants and plant products due to infection of the plant. Plant diseases, as a result of infection by plant pathogens, cause damage to the plants and/or plant products, reduce production and yield, limit the kind of plants that can grow in certain geographic areas and as a result cause severe (financial) losses to the grower.

Different means for control of the plant pathogens exist, such as mechanical cultivation of the soil, chemical treatment with pesticides, including fungicides and insecticides, or crop rotation. However, for certain plant pathogens, especially soil born fungi, these means of control are insufficient to protect the plants from infection and resulting diseases. The only effective means of control involves plant host resistance (Russell, 1978, Plant Breeding for pest and disease resistance, Butterworths edit., 485 pp). The development of cultivars resistant to common plant pathogens is one of the major goals of plant breeders today, in order to reduce or ultimately eliminate the extensive need for pesticides. The burden for the environment of the large amounts of pesticides sprayed on crops, trees etc. worldwide each year becomes too severe. Moreover, governmental regulations in Western countries restrict the use or even forbid the use of certain pesticides. Therefore, the need for plants which are resistant to one or more of their pathogens, or which have a reduced susceptibility to their attackers becomes more and more pressing.

Plants have developed a complex defense mechanism against attack and infection by pathogens. In general, their defense system is twofold, at the one hand it comprises a general resistance which is effective against different pathogen species, and at the other hand, it consists of a strong resistance against specific pathogen species. This latter resistance is generally based on a hypersensitivity reaction (HR). Although the exact defense mechanisms of the plant have still to be elucidated, it is assumed that when a pathogen comes into contact with a host cell, an early event takes place that triggers a rapid response that impedes further growth of the pathogen and subsequent development of the disease. Frequently within the pathogen population genotypes exist which can overcome this HR-resistance, so called races. If a host is resistant to a specific pathogen race, but susceptible to another race, one then speaks of race-specific resistance. It is furthermore believed that such a defense system in plant-pathogen interactions is based on a gene-for-gene relationship (Flor, 1956, Adv.Gen. **8**, 29-54). In the gene-for-gene model it is postulated that for each gene conferring resistance to the host, there is a corresponding gene in the pathogen that confers avirulence, and vice versa. Very recently, evidence on molecular level for the gene-for-gene concept has been found with the cloning of a few plant resistance genes. The resistance gene of a plant encodes for a product, a receptor molecule, that can recognize a product of the pathogen, an elicitor, encoded by a avirulence gene. If the receptor interacts with the elicitor molecule a hypersensitive response is triggered, resulting into the destruction of the infected plant cells and surrounding cells, and so preventing the multiplication and spread of the pathogen within the plant. This postulated mechanism has recently been confirmed by the isolation of a few avirulence genes of the pathogen with the corresponding resistance genes of the host plant. Examples of such resistance genes are: *RPS2* from Arabidopsis (resistance to *Pseudomonas syringae* expressing avrRpt2), *N* from tobacco (resistance to tobacco mosaic virus), *Cf-9* from tomato (resistance to the leaf fungal pathogen *Cladosporium fulvum* carrying *avr9*) and L^6 from flax (resistance to the corresponding leaf rust fungal race) (Dangl, 1995, Cell **80**, 363-366).

As stipulated before, the development of resistant plants is one of the important objectives of current plant breeding programs. Plant genotypes susceptible for particular pathogens are crossed with resistant plant genotypes in order to introduce the resistant phenotype into the breeding line. However, the breeding of resistance genes is restricted due to several factors: (i) the limited occurrence of (known) resistance genes in the available germplasm, (ii) incompatibility of crossing between different species and (iii) the limited availability of reproducible and reliable disease assays for certain pathogens, said assays being a prerequisite in selection breeding.

Amongst the pathogens causing serious damages to plants, one finds the group of soilborne cortical rots and vascular wilt inducing fungi, such as Fusarium and Verticillium (Toussoun, 1981, in: Fusarium Diseases, Biology and Taxonomy, Nelson, Toussoun & Cook edit., Penn. State Univ. Press, 457 pp.). Said wilt inducing fungi infect the plants through the roots via direct penetration or via wounds after which the xylem vascular tissue of the plant is colonized and symptoms of infection with said fungi are wilting, browning and dying of leaves followed by plant death. Entire plants or plant parts above the point of vascular invasion of the pathogen may die within a period of some weeks after infection. The fungi usually spread internally through the xylem vessels as mycelium or conidia until the entire plant is killed. Because of the fact that those fungi are able to survive in the soil saprophytically, they become established forever once they are introduced in the field. They are distributed more or less worldwide causing tremendous losses on most species of vegetables and flowers, field crops, fruit trees, etc. Because those fungi are so widespread and so persistent in soils the only effective way of controlling said wilt inducing fungi is using resistant plant genotypes.

Most Fusarium species belong to the family of imperfect fungi. This class of fungi is characterized by the fact that only a vegetative stage of the fungus is known. The generative stage of those fungi has not been discovered yet. Due to the overall classification or taxonomy of fungi upon their morphological characteristics of the generative phase, one should bear in mind that fungi belonging to the class of imperfect fungi can be classified into another class once

their generative stage is discovered and, subsequently, a change of classification and name can follow (Gerlach, 1981, in: Fusarium, Diseases, Biology and Taxonomy, Nelson, Toussoun & Cook edit., Penn. State Univ. Press, p. 413-426). Moreover, it has been observed that some Fusarium species can "mutate" into another species depending on the plant infected and/or the environment (Bolton & Davidson, 1972, Can. J. Plant Sci, **52**, 189-196). Up to now most of the wilt inducing Fusarium belong to the species *Fusarium oxysporum*. Different plant species are attacked by different races or isolates of Fusarium (Armstrong & Armstrong, 1981, in: Fusarium, Diseases, Biology and Taxonomy, Nelson, Toussoun & Cook edit., Penn. State Univ. Press, p. 391-399). However, some Fusarium isolates can infect different plant species. Known Fusarium isolates are for example: *Fusarium oxysporum* f.sp. *lycopersici* (tomato), *F. oxysporum* f.sp. *melonis* (melon), *F. oxysporum* f.sp. *batatas* (sweet potato, tobacco), *F. oxysporum* f.sp. *cepae* (onion), *F. oxysporum* f.sp. *conglutinans* (cabbage, radish), *F. oxysporum* f.sp. *cubense* (banana), *F. oxysporum* f.sp. *vasinfectum* (cotton, alfalfa, soybean, tobacco) *F. oxysporum* f.sp. *dianthii* (carnation), *F. oxysporum* f.sp. *chrysanthemi* (chrysanthemum), *F. oxysporum* f.sp. *tuberosi* (potato), *F. oxysporum* f.sp. *cyclaminis* (cyclamen), *F. oxysporum* f.sp. *nicotianae* (tobacco).

In the light of the present invention it should be recognized that the name of the wilt inducing fungi can change in the future, but this will not effect the scope of the invention.

The isolation of plant genes without knowing their gene products is like looking for a needle in a haystack, because of the enormous genome sizes of plant species: e.g. tomato has a genome size of 1000 Mb (10^9 base pairs of nuclear DNA), maize has a genome size of 3000 Mb and wheat has even more than 16 x 10^9 base pairs. Searching for a specific gene among these billions of base pairs is only feasible when (i) there are enough molecular markers tightly linked to the gene of interest and (ii) there is good genetic material available (Tanksley *et al.*, 1995, Trends in Genetics, **11**, p. 63-68).

SUMMARY OF THE INVENTION

The present invention relates to a nucleic acid comprising the *I-2* resistance gene which when present and expressed in a plant is capable of conferring said plant resistant against wilt inducing fungi. Furthermore, the invention relates to the *I-2* resistance gene of which the DNA sequence is disclosed herein. The invention also relates to a gene product encoded by the *I-2* resistance gene which is capable of triggering a hypersensitive response in the plant when it comes into contact with a gene product encoded by a corresponding avirulence gene of the plant pathogen. In addition the present invention relates to DNA constructs, cosmids, vectors, bacterial strains, yeast cells and plant cells comprising the *I-2* resistance gene. In another aspect, the present invention relates to a genetically transformed plant, which is resistant to a wilt inducing fungus, said fungus being capable of infecting the untransformed plant. Furthermore, the invention relates to resistance genes which are homologous to the *I-2* resistance gene, and which, when present in a plant, are able of conferring said plant resistant to infection by pathogens.

Finally, the invention relates to oligonucleotides corresponding to the sequence of the *I-2* resistance gene or part thereof, and detection kits comprising said oligonucleotides.

DESCRIPTION OF THE FIGURES

Figure 1 shows a schematic representation of YAC 1/546, with a size of 750 kb, and the position of the *Bss*HII, *Rsr*I, *Sfi*I and *Sgr*A1 restriction sites (indicated by small lines). The hatched bar represents the 255 kb *Sgr*A1 fragment comprising the *I-2* resistance gene. The most lower line represents the size bar (in kb). The circle/arrowhead combination represents the left arm of pYAC4, direction of the centromer.

Figure 2 shows a schematic drawing of the binary cosmid vector pJJ04541 which is used to construct a cosmid library of YAC 1/546. Plasmid pRK290 (20 kb large) (Ditta *et al*, 1980, Proc. Natl. Acad. Sci. USA, **77**, 7347-7351) was used as starting vector."Tet" refers to the gene conferring resistance to tetracyclin. "LB" signifies T-

DNA left border repeat sequence, and "RB" signifies the right border repeat. The cauliflower mosaic virus 35S promoter sequence is indicated by "p35S", and "ocs3"' indicates the octopine synthase 3' end. "NPT" indicates neomycin phosphotransferase, and "cos" refers to the bacteriophage lambda cos site enabling *in vitro* packaging. "pDBS" indicates the polylinker of pBluescript (Stratagene, La Jolla, CA, USA).

Figure 3 shows part of a 4,5% denaturing polyacrylamide gel with DNA fingerprints of 24 cosmids using Restriction Fragment Amplification with the enzyme combination *Eco*RI/*Mse*I. The templates used are depicted in the right part of the figure.

Figure 4 shows a schematic representation of the 255 kb *Sgr*A1 fragment with the position of the *Mlu*I and *Sal*I restriction sites (indicated with small lines) and the position of the 18 AFLP markers EM01 to EM18 (indicated with arrows). The cosmid contig of the DNA segment flanked by markers 18 and 12 is indicated with horizontal lines. The cosmids marked with an asterisk are used in the complementation analysis. The upper line represents the size bar (in kb) and the sizes of the *Mlu*I and *Sal*I restriction fragments are indicated.

Figure 5 shows a schematic representation of the overlapping cosmids A52 and B22 and partly of A55 and CC16 (the open arrow head indicates that the cosmid is continued). The position of the various restriction sites is indicated with small lines. The position of the AFLP markers EM05, EM14 and EM06 is indicated with an arrow. The DNA segment of which the nucleotide sequence was determined is indicated with a line with a bidirectional arrow.

Figure 6 shows the nucleotide sequence of a DNA segment of almost the complete overlap between cosmids A52, B22 and A55, and the deduced amino acid sequence of the *I-2* resistance gene. The initiation codon (ATG position 1798-1800) is underlined and the termination codon (TAA position 5596-5598) is double underlined.

The position of the AFLP marker EM06 is from nucleotide position 3470 (5'-AATTC**AGA**-3') to nucleotide position 3565 (5'-**AGA**TTA-3').

The positions of three intron sequences are given in italics: one intron of 86 nucleotides located upstream of the ATG initiation codon from nucleotide position 1703 to 1788 and two introns of respectively 399 and 82 nucleotides located downstream of the TAA termination codon from respectively nucleotide position 5628 to 6026 and 6093 to 6174. The transcriptional initiation site is predicted to be located at least 201 nucleotides upstream of the ATG initiation codon. The transcriptional termination site is predicted to be located at least 893 nucleotides downstream of the TAA termination codon. A putative poly-adenylation signal (AAUAAA) is located at nucleotide position 6406-6411 and is given in bold.

Figure 7 shows a schematic drawing of plasmid pKG6016. "sm/sp adtr" refers to the streptomycin/spectinomycin resistance gene, the origin of replication of pBR322 is indicated by "OriV", and "bla" refers to the ampicilin resistance gene. "LB" signifies T-DNA left border repeat sequence, and "RB" signifies the right border repeat. The nopaline synthase promoter sequence is indicated by "nos pr", and "nos 3'" indicates the nopaline synthase 3' end. "nptII" indicates the kanamycin resistance gene. "I2-upstream" refers to the 1.3 kb DNA segment upstream of the coding sequence of the *I-2* resistance gene, "FusI2" refers to the coding sequence of the *I-2* resistance gene from nucleotide position 1798 to nucleotide position 5598, and "3' untrans" refers to the 1.1 kb DNA segment downstream of the coding sequence of the *I-2* resistance gene. The relevant restriction sites are indicated.

DETAILED DESCRIPTION OF THE INVENTION

In the description and examples that follow, a number of terms are used herein. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

- nucleic acid: a double-stranded DNA molecule;
- oligonucleotide: a short single-stranded DNA molecule;

- primers: in general, the term primer refers to a single-stranded DNA molecule which can prime the synthesis of DNA;
- nucleic acid hybridization: a method for detecting related DNA sequences by hybridization of single-stranded DNA on supports such as nylon membrane or nitrocellulose filter papers. Nucleic acid molecules that have complementary base sequences will reform the double-stranded structure if mixed in solution under the proper conditions. The double-stranded structure will be formed between two complementary single-stranded nucleic acids even if one is immobilized on a support. In a Southern hybridization procedure, the latter situation occurs;
- hybridization probe: to detect a particular DNA sequence in the Southern hybridization procedure, a labelled DNA molecule or hybridization probe is reacted to the fractionated DNA bound to a support such as nylon membrane or nitrocellulose filter paper. The areas on the filter that carry DNA sequences complementary to the labelled DNA probe become labelled themselves as a consequence of the reannealing reaction. The areas of the filter that exhibit such labelling can then be detected according to the type of label used. The hybridization probe is generally produced by molecular cloning of a specific DNA sequence or by synthesizing a synthetic oligonucleotide;
- homologous sequence: a sequence which can hybridize under stringent conditions to a particular sequence, and/or a DNA sequence coding for a polypeptide which has the same properties as the polypeptide encoded by the particular DNA sequence, and/or a DNA sequence coding for a polypeptide having the same amino acid sequence as the polypeptide encoded by the particular DNA sequence and/or an amino acid sequence in which some amino acid residues have been changed with respect to the amino acid sequence of the particular polypeptide without substantial effect on the major properties of said polypeptide and/or a sequence which has at least 50 %, preferably 60 %, more preferably 70 %, most preferably 80 % or even 90 % sequence identity with the particular sequence, whereby the length of sequences to be compared for nucleic acids is generally at least

120 nucleotides, preferably 200 nucleotides and more preferably 300 nucleotides and the length of sequences to be compared for polypeptides is generally at least 40 amino acid residues, preferably 65 amino acid residues and more preferably 100 amino acid residues;

- promoter: a transcription regulation region upstream from the coding sequence containing the regulatory sequences required for the transcription of the adjacent coding sequence and includes the 5' non-translated region or so called leader sequence of mRNA;
- terminator: a region downstream of the coding sequence which directs the termination of the transcription, also called the 3' non-translated region, which includes the poly-adenylation signal;
- resistance gene: a nucleic acid having a coding sequence as depicted in Figure 6, or part thereof, or any corresponding or homologous coding sequence;
- stringent conditions refer to hybridization conditions which allow a nucleic acid sequence to hybridize to a particular sequence. In general, high stringent conditions refer to the hybridization conditions which allow a nucleic acid sequence of at least 50 nucleotides and preferably about 200 or more nucleotides to hybridize to a particular sequence at about 65 °C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at 65 °C in a solution comprising about 0,1 M salt, or less, preferably 0,2 x SSC or any other solution having a comparable ionic strength. These conditions allow the detection of sequences having about 90 % or more sequence identity. In general, lower stringent conditions refer to the hybridization conditions which allow a nucleic acid sequence of at least 50 nucleotides and preferably about 200 or more nucleotides to hybridize to a particular sequence at about 45 °C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at room temperature in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic

strength. These conditions allow the detection of sequences having up to 50 % sequence identity. The person skilled in the art will be able to modify these hybridization conditions in order to identify sequences varying in identity between 50 % and 90 %.

Alternatively, stringent conditions refer to hybridization conditions which allow a nucleic acid sequence to hybridize selectively to the *I-2* resistance gene in its genomic environment, substantially to the exclusion of hybridization with other DNA sequences of said genomic environment.

- Fusarium 2: *Fusarium oxysporum* f.sp. *lycopersici* race 2 or any other genotype which is not able to infect a host having a resistance gene according to the invention; other genotypes are such as but not limited to, wilt inducing fungi, soil born fungi, or any other plant pathogens.
- resistance gene product: a polypeptide having an amino acid sequence as depicted in Figure 6, or part thereof, or any homologous amino acid sequence;
- R_0 plant: primary regenerant from a transformation experiment, also denoted as transformed plant or transgenic plant;
- R_1 line: the progeny of a selfed R_0 plant.
- R_2 line: the progeny of a selfed R_1 plant.
- R_1BC line: the progeny of a backcross between a R_1 plant and a plant of the genotype which was originally used for the transformation experiment.

In the present invention we have been able to identify and isolate the *Immunity-2* (*I-2*) resistance gene. The gene was cloned from a tomato genotype which is resistant to *Fusarium oxysporum* f.sp. *lycopersici* race 2. The isolated *I-2* resistance gene according to the invention can be transferred to a susceptible host plant using Agrobacterium mediated transformation or any other known transformation method, and is able to confer the host plant resistant against Fusarium 2. The host plant can be tomato or any other genotype that is infected by Fusarium 2.

The present invention provides also the nucleic acid sequence of the *I-2* resistance gene which is depicted in Figure 6.

With the *I-2* resistance gene according to the invention, one has an effective means for control against wilt inducing fungi, since the gene can be used for transforming susceptible plant genotypes thereby producing genetically transformed plants having a reduced susceptibility or being preferably resistant to infection by wilt inducing fungi. In a preferred embodiment the *I-2* resistance gene comprises the coding sequence preceded by a promoter region and followed by a terminator region. The promoter region should be functional in plant cells and, preferably, corresponds to the native promoter region of the *I-2* resistance gene. However, it should be recognized that any heterologous promoter region can be used in conjunction with the coding sequences, as long as it is functional in plant cells. Preferably, a constitutive promoter is used, such as the CaMV 35 S promoter or T-DNA promoters, all well known to those skilled in the art. Furthermore, a suitable terminator region should be functional in plant cells all well known to those skilled in the art.

In addition the invention relates to the *I-2* resistance gene product which is encoded by the *I-2* resistance gene according to the invention and which has an amino acid sequence provided in Figure 6, or which is homologous to the deduced amino acid sequence or part thereof as listed in Figure 6. The *I-2* resistance gene product can be used for the identification and/or isolation of the corresponding gene product encoded by an avirulence gene of the pathogen. The relationship between the *I-2* resistance gene product, which is assumed to be acting like a receptor molecule, and the gene product of the pathogen, which is assumed to be acting like an elicitor molecule, is characterized by the occurrence of a defense mechanism reaction in the plant. Furthermore the *I-2* resistance gene product can be used for raising antibodies against it, which antibodies can be used for the detection of the presence of the *I-2* resistance gene product.

In another aspect of the invention, the *I-2* resistance gene can be used for the design of oligonucleotides which are complementary to one strand of the DNA sequence as described in Figure 6, or part thereof, which can be used as hybridization probes, being accordingly labelled to allow detection, for the screening of genomic DNA or cDNA libraries for homologous genes. Homologous sequences which can hybridize to the probe, and which encode for a gene product

that is able to confer resistance to a plant against a fungus which normally infects said plant, or both are comprised within the scope of the present invention.

In another aspect of the invention oligonucleotides are designed based on the *I-2* resistance gene sequence, such that they can be used as hybridization probes in Southern analysis. These probes can be used as molecular markers to distinguish plant genotypes having the resistance gene and plant genotypes lacking the resistance gene. Such a probe can be used as an additional tool in selection breeding. In a preferred embodiment of the invention, oligonucleotides are designed based on the *I-2* resistance gene sequence, such that they can be used as primers in an amplification reaction, such as polymerase chain reaction (PCR), whereby the formation of an amplification product indicates the presence of the *I-2* resistance gene in a certain plant genotype. In a particular embodiment of the invention said primers direct the amplification of polymorphic fragments, so called molecular markers, which are closely linked to the *I-2* resistance gene. The invention also relates to diagnostic kits, comprising oligonucleotides according to the invention, for the detection of the presence or absence of the *I-2* resistance gene within a genotype under study. Such a diagnostic kit circumvents the use of a laborious disease assay to screen for genotypes having the resistance gene or not.

Furthermore the invention relates to DNA constructs (A) comprising a DNA sequence corresponding to the coding sequence of the *I-2* resistance gene and regulatory sequences functional in plant cells. Said regulatory sequences are either homologous or heterologous to the coding sequences of the *I-2* resistance gene. The invention relates also to DNA constructs (B) comprising the regulatory sequences, and more preferably the promoter region of the *I-2* resistance gene in conjunction with a structural gene sequence heterologous to said regulatory sequences.

The invention relates also to a DNA vector comprising a DNA construct (A) and/or a DNA construct (B). Suitable vectors can be cloning vectors, transformation vectors, expression vectors, etc...., which are well known to the person skilled in the art. Furthermore, cells harbouring a vector comprising a DNA

sequence corresponding to the sequence as described in Figure 6 or part thereof, DNA constructs (A) or DNA constructs (B), are within the scope of the invention.

In one preferred embodiment of the invention, a genetically transformed plant is obtained by introducing the *I-2* resistance gene within the genome of said plant, having a susceptible genotype to Fusarium 2, using standard transformation techniques, wherein said genetically transformed plant is resistant to Fusarium 2. Moreover, the *I-2* resistance gene is inherited in following generations of the said genetically transformed plant and is able to confer the next generations of said plant resistant to Fusarium 2.

In yet another embodiment of the invention part of the DNA sequence comprising the *I-2* resistance gene, is used for transforming a plant which is susceptible to Fusarium 2. Such part can be obtained by digesting the DNA sequence comprising the *I-2* resistance gene, in one or more steps, with one or more appropriate restriction enzymes, chosen on the basis of the presence of their recognition site in the *I-2* resistance gene according to the invention, or in the sequences flanking the *I-2* resistance gene. The obtained DNA segment can be transferred to a susceptible host plant and genetically transformed plants having a resistant phenotype can be identified when inoculated with Fusarium 2.

We have found that the *I-2* resistance gene according to the present invention, is functional in homologous systems and/or heterologous systems, such as but not limited to tomato, melon, tobacco, Arabidopsis, egg plant, potato species, and is involved in reducing the susceptibility and/or is capable of conferring these plant species resistance against *Fusarium 2* as defined above, and especially against one or more wilt inducing fungi. A homologous system refers to a plant species which is the same plants species from which the resistance gene was isolated and a heterologous system refers to a plant species which is different from the plant species from which the resistance gene was isolated.

The DNA sequence comprising the *I-2* resistance gene as provided in the present invention has numerous applications of which some are described herein but which are not limiting the scope of the invention.

The present invention will be further described in detail in view of the isolation of the *I-2* resistance gene present in tomato lines which are resistant against *Fusarium oxysporum* f.sp. *lycopersici* race 2. For the isolation of the *I-2* resistance gene we have used a map-based cloning (positional cloning) strategy, comprising the following steps:

(1) identification of molecular markers linked to the *I-2* resistance gene,

(2) genetic mapping of the *I-2* locus using morphological markers,

(3) construction of a high molecular weight genomic YAC library,

(4) physical mapping of the molecular markers on the YAC clones and YAC contig building,

(5) construction of a cosmid library of the YAC clone harbouring the linked molecular markers,

(6) physical fine mapping and cosmid contig building,

(7) transformation of susceptible plants with the cosmids forming the contig,

(8) complementation analysis.

For the identification of molecular markers, we have used the selective restriction fragment amplification technology, hereinafter denoted as AFLP™ technology, which randomly amplifies a subset of DNA fragments out of a complex mixture of many DNA fragments and said amplified fragments generate fingerprints that can be analyzed.

In general, total DNA of different genotypes of the same plant species are subjected to the AFLP technology and the different AFLP fingerprints obtained from the different genotypes are compared. Fragments that are present in one genotype and absent in another genotype are polymorphic fragments and are denoted as AFLP markers.

The selectivity in AFLP reactions is obtained by using randomly chosen selective nucleotides at the 3' end of the PCR primers immediately adjacent to the nucleotides of the restriction enzyme site. In an AFLP screening the DNA to be studied is subjected to different primer combinations. The total amount of different primers that can be used is determined by the number of selective nucleotides that are added to the 3' end (4 primers with 1 selective nucleotides, 16 primers with 2 selective nucleotides, 64 primers with 3 selective nucleotides). If two different

restriction enzymes are used than there are twice the amount of primers. Those primers can be used in different combination. If all possible combinations are used in an AFLP screening, than all the fragments present should have been amplified with one of the primer combinations (Zabeau and Vos, EP 0534858).

For the identification of AFLP markers linked to the *I-2* resistance gene different tomato lines were subjected to an AFLP screening. The tomato lines were pooled into different pools: on the one hand pools of tomato lines being resistant against Fusarium 2 and on the other hand pools of tomato lines being susceptible to Fusarium 2.

The pools were subjected to an AFLP screening using the *Eco*RI/*Mse*I enzyme combination.

The following primers are used for the AFLP screening:

*Eco*RI-primers: 5'-GACTGCGTACCAATTCNNN-3'

*Mse*I-primers: 5'-GATGAGTCCTGAGTAANNN-3'

The N's indicate the variable selective nucleotides. In the AFLP screening all 64 possible primers were used for both the *Eco*RI- and *Mse*I-primer giving a total of 64 x 64 = 4096 primer combinations. The objective of the screening was to identify AFLP markers linked to the *I-2* resistance gene, i.e. present in the fingerprints of the resistant pools, and absent in the fingerprints of susceptible pools. In the analysis of all the AFLP fingerprints a total of 18 AFLP markers were identified which were present in the resistant pools and absent in the sensitive pools: these markers are referred to as candidate *I-2* linked AFLP markers.

The presence of these candidate *I-2* linked AFLP markers were confirmed on the individual tomato lines. All markers were present in the resistant lines and absent in the susceptible lines.

However, the position of the AFLP markers with respect to the *I-2* resistance gene and the distance between the *I-2* resistance gene and the respective markers has still to be determined. For that purpose a genetic map of the *I-2* locus was made by making genetic crosses between Fusarium 2 resistant tomato lines and susceptible tomato lines having morphological markers, followed by a screening for recombinants in the segregating F2 populations and screening of the F2 recombinants with the AFLP markers. The results indicated that the *I-2*

locus is flanked by marker EM18 at one end of the DNA segment comprising the *I-2* resistance gene and markers EM03, EM12 and EM16 at the other end. All the other AFLP markers co-segregated with the *I-2* resistance gene.

Next, the AFLP markers were screened on a high molecular weight genomic library. The cloning of very large segments of DNA as large artificial chromosomes in yeast has become an essential step in isolating genes via positional cloning. The cloning capacity of the YAC vector allows the isolation of DNA fragments up to one million base pairs in length. The tomato line *Lycopersicon esculentum* E22, homozygous for the *I-2* locus, was used as source DNA to construct a YAC library. We obtained a YAC library containing 3840 clones with an average insert size of 520 Kb, representing approximately 2.2 genome equivalents of the tomato genome. One positive clone was obtained after an AFLP screening with the *I-2* linked AFLP markers, and all markers were present on this individual YAC clone, designated as YAC 1/546. The size of this YAC clone is determined to be 750 kb.

Further analysis has determined that all of the AFLP markers were located on a *Sgr*AI fragment, at a distance of 255 kb of the left arm of the YAC until the first *Sgr*AI site. A schematic representation of the physical map of the left arm of YAC 1/546, containing the *I-2* locus, including the location of the 18 AFLP markers, is depicted in Figure 4. The exact location of the AFLP markers can be determined on the basis of mapping cosmids on the physical map.

The size of an insert in YAC 1/546 is still too large for direct localization of the *I-2* gene. Such large inserts cannot be transformed into plant cells directly. Therefore, a cosmid library was constructed of the yeast strain containing YAC 1/546 using cosmid vectors which are suitable for *Agrobacterium* mediated transformation. The size of this binary cosmid vector amounts 29 kb and is shown schematically in Figure 2. The cloning capacity of this binary cosmid vector, using phage lambda packaging extract is within the range of 9 to 24 kb. A bank of approximately 250,000 cosmid clones was obtained from size fractionated yeast DNA. The cosmid bank was screened by colony hybridization using the labelled *Sgr*AI fragment as probe. Of about 10,000 colonies approximately 150 positive cosmid clones were identified.

In the following step, the position of the AFLP markers on the 255 kb SgrAI fragment was determined on the basis of a cosmid contig. The positive cosmid clones were screened with the 18 AFLP markers and their position was determined. A schematic outline of the cosmid contig and the physical fine mapping of the 18 AFLP markers is depicted in Figure 4.

The final step in the identification of the *I-2* resistance gene via positional cloning is the complementation of the corresponding susceptible phenotype. The cosmid clones were introduced in *Agrobacterium tumefaciens* through conjugative transfer in a tri-parental mating. The presence of the cosmid in the *A. tumefaciens* strains was determined comparing various restriction enzyme patterns as well as DNA fingerprints from the *A. tumefaciens* strains with the *Escherichia coli* strain containing the cosmid. Only those *A. tumefaciens* cultures harbouring a cosmid with the same DNA pattern as the corresponding *E. coli* culture were used to transform a susceptible tomato line.

A susceptible tomato line was transformed with several cosmids forming the cosmid contig using standard transformation methods.

The primary regenerants (R_0 plants) of the transformation experiments were grown in the greenhouse for seed set to obtain R_1 lines. These were tested for disease symptoms in order to identify cosmids with the resistance gene. The disease assay is performed on seedlings. Their roots are immersed in a conidial suspension of *Fusarium oxysporum* f.sp. *lycopersici* race 2 and disease symptoms are scored three to four weeks after inoculation. Plants are scored resistant when they are healthy without wilting symptoms and/or without browning of the stem tissue. Plants being dead or having yellow wilting leaves and having severe browning of the stem tissue are scored susceptible. The observations of the disease assay revealed that 3 cosmids were able to complement the susceptible phenotype, thereby providing definitive evidence that a functional *I-2* resistance gene is located on each of these 3 cosmids.

To demonstrate that the resistant phenotype of the transgenic R_0 plants, which were transformed with the overlapping cosmids A52, B22 and A55, is determined by the genomic insert present in the various cosmids, the presence of the corresponding AFLP markers was investigated (see Figure 5). Selective

restriction fragment amplification was performed with the primer combinations identifying the markers EM14 and EM06 for the R_0 plants transformed with cosmids A52 or B22 and with the primer combinations identifying the markers EM06 and EM04 for the R_0 plants transformed with cosmid A55. The DNA fingerprints obtained showed in both cases and for both markers the presence of the markers in the resistant plants and the absence of the markers in the susceptible plants indicating that the three identified overlapping cosmids A52, B22 and A55 comprise the *I-2* resistance gene.

In order to confirm the stable integration of the *I-2* resistance gene into the genome of the transgenic R_0 plants, resistant plants of the R_1 lines were selfed and grown in the greenhouse for seed set to obtain R_2 lines. Seedlings of the R_2 lines were subjected to a disease assay as described above and were scored for disease symptoms: Wilting plants were considered to be susceptible, whereas plants showing no wilting were considered to be resistant. The results obtained indicated the stable Mendelian inheritance of the *I-2* resistance gene. Additionally, resistant plants of the R_1 lines were backcrossed with the susceptible tomato genotype used for the transformation experiments to obtain R_1BC lines. The results of the disease assay performed on seedlings of the R_1BC lines confirmed both the inheritance as well as the dominance of the *I-2* resistance gene.

Finally, the inserts in cosmids A52, B22 and A55 were further characterized, and the minimal DNA segment comprising the *I-2* resistance gene, defined by the left end of cosmid A55 and the right end of cosmid B22, was sequenced. Sequencing analysis revealed a large open reading frame of 3798 nucleotides. The DNA sequence is listed in Figure 6.

The DNA sequence comprising the *I-2* resistance gene was further subjected to transcript mapping studies in order to identify the transcription initiation site and transcription termination site and to determine the existence of intron sequences. These transcript mapping studies were performed according to generally known methods whereby genomic DNA sequences are compared with cDNA sequences. The comparison of cDNA sequences and genomic sequences revealed the existence of three intron sequences in the *I-2* resistance gene outside the coding sequence: one intron of 86 nucleotides is located upstream of the ATG

initiation codon and two introns of 399 and 82 nucleotides respectively are located downstream of the TAA termination codon, as is depicted in Figure 6. The transcription initiation site is located at or upstream of nucleotide 1597. The transcription termination site is located at or immediatly downstream of nucleotide 6491. The position of a putative poly-adenylation signal is deduced at nucleotides 6406-6411 (AAUAAA).

A DNA segment corresponding to part of the DNA sequence, as provided in Figure 6, starting at nucleotide position 464 and ending at nucleotide position 6658 was used for transforming a susceptible tomato genotype. The DNA segment was obtained by digesting cosmid B22 with restriction enzymes *Bam*HI and *Sal*I, providing a 3.8 kb fragment, and with *Sca*I and *Bam*HI, providing a 2.4 kb fragment, resulting in a 6.2 kb fragment comprising the coding sequence of the *I-2* resistance gene flanked upstream by a 1.3 kb DNA sequence and downstream by a 1.1 kb DNA sequence. The DNA segment was cloned into a suitable cointegrate type vector and subsequently introduced through *Agrobacterium tumefaciens* mediated transformation into a tomato plant which is susceptible to Fusarium 2. The R_0 plants were grown in the greenhouse for seed set to obtain R_1 lines and these R_1 lines were subjected to the disease assay as described above. The observations indicated that the DNA segment is involved in conferring to the transformed plants a reduced susceptibility to Fusarium 2.

It is understood that those skilled in the art can choose other parts of the DNA sequence on the basis of the teaching provided herewith, or using any other well known method, to obtain parts of the DNA sequence provided in Figure 6, or to obtain parts of the genomic insert present in cosmid B22 or A55, introducing said parts into an appropriate DNA construct which allow expression in plant cells, transferring said constructs to a host plant being susceptible to Fusarium 2 and testing the transformed plants for disease symptoms after inoculation with Fusarium 2.

Cosmid B22 was used for the transformation of susceptible genotypes of melon, tobacco as well as Arabidopsis according to general known transformation methods. The R_0 plants were grown in the greenhouse for seed set to obtain R_1

lines. These were tested for disease symptoms in order to identify the functionality of the *I-2* resistance gene. The disease assay was performed on seedlings as described herein. Cosmid B22 was also used for the transformation of a susceptible genotype of potato. Vegetatively propagated transformed plants were obtained and were subjected to a disease assay on cuttings. The observations of the disease assay on the transformed plants revealed the complementation of the susceptible phenotype.

For the identification and isolation of homologous sequence falling within the scope of the present invention, genomic and cDNA libraries were screened with the coding sequence of the *I-2* resistance gene as a probe under stringent hybridization conditions. Positive clones were isolated and were used for complementation analysis.

Cosmid B22 has been deposited on July 14, 1995 as plasmid pKGI2-B22 at Centraalbureau voor Schimmelcultures at Baarn, The Netherlands, under deposit number CBS 546.95.

Cosmid A55 has been deposited on August 5, 1996 as plasmid pKGI2-A55 at Centraalbureau voor Schimmelcultures at Baarn, The Netherlands, under deposit number CBS 820.96.

The following examples will provide a further illustration of the present invention which is nevertheless not limited to these examples.

EXAMPLES

EXAMPLE 1: DISEASE ASSAY

Fusarium oxysporum forma specialis *lycopersici* race 2 was maintained on Czapek Dox Agar (Difco Laboratories, Detroit, MI, USA). Conidial suspensions were obtained by culturing the fungus in Czapek Dox Broth (Difco Laboratories, Detroit, MI, USA) on a reciprocal shaker for 4 to 7 days at 25° Celsius. The conidia were separated from mycelium fragments by filtration through a stainless steel filter with a pore size of 50 μm. The suspensions were adjusted to a concentration of 2×10^6 conidia per ml by diluting with water.

Seedlings of tomato

Seeds of tomato were germinated in soil in the greenhouse at 25° Celsius. Ten to 14 days-old seedlings were used for inoculation with the fungus. The seedlings were carefully pulled out of the soil and the roots were dipped in water for removing most of the adhering soil. Subsequently, the roots were immersed in the conidial suspension for two minutes and the plants were repotted in soil. The plantlets were grown in the greenhouse at a temperature of 25° C at daytime (16 hours) and 22° C at night (8 hours). After three to four weeks the plants were scored for disease symptoms.

The plants were evaluated as follows: Resistant plants resemble non-inoculated control plants; they are large non-wilting and/or without browning of stem tissue. Susceptible plants are dead or show typical symptoms: small plants with yellow, wilting leaves and severe browning of stem tissue.

Cuttings of tomato

Cuttings were prepared from greenhouse-grown R_0 plants. Small sideshoots were cut from the plants and put in soil under 100% humidity at 20° Celsius. After one to two weeks the cuttings started rooting. Two to three weeks old cuttings were used for inoculation with the fungus. The plantlets were carefully pulled out of the soil. The roots were dipped in water for removing most of the adhering soil. Subsequently, the roots were immersed in the conidial suspension for five minutes and the plants were repotted in soil. The plantlets were grown in the greenhouse at a temperature of 25 °C at daytime (16 hours) and 22 °C at night (8 hours). After three to four weeks the plants were scored for disease symptoms.

Evaluation of the plants was as described for seedlings of tomato.

EXAMPLE 2: IDENTIFICATION OF AFLP MARKERS LINKED TO A DNA SEGMENT COMPRISING THE *I-2* RESISTANCE GENE

Tomato lines (*Lycopersicon esculentum*)

A total of 10 *F. oxysporum* f.sp. *lycopersici* race 2 resistant and 10 *F. oxysporum* f.sp. *lycopersici* race 2 susceptible tomato lines were used, and are depicted below:

1.	DR9	resistant	De Ruiter Zonen C.V., Bleiswijk, The Netherlands (hereinafter "De Ruiter")
2.	RZ5	resistant	Rijk Zwaan Zaadteelt en Zaadhandel B.V., De Lier, The Netherlands (hereinafter "Rijk Zwaan")
3.	E22	resistant	Enza Zaden, de Enkhuizer Zaadhandel B.V., Enkhuizen, The Netherlands (hereinafter "Enza Zaden")
4.	E16	resistant	Enza Zaden
5.	DR4	resistant	De Ruiter
6.	RZ7	resistant	Rijk Zwaan
7.	E3	resistant	Enza Zaden
8.	E7	resistant	Enza Zaden
9.	RZ4	resistant	Rijk Zwaan
10.	DR12	resistant	De Ruiter
11.	GCR210	susceptible	Institute of Horticultural Research, Littlehampton, Great Britain
12.	GCR508	susceptible	Institute of Horticultural Research, Littlehampton, Great Britain
13.	52201	susceptible	Rijk Zwaan
14.	DR5	susceptible	De Ruiter
15.	E12	susceptible	Enza Zaden
16.	E1	susceptible	Enza Zaden
17.	RZ8	susceptible	Rijk Zwaan
18.	E6	susceptible	Enza Zaden
19.	RZ10	susceptible	Rijk Zwaan
20.	DR11	susceptible	De Ruiter

Isolation and modification of the DNA

Total tomato DNA from the 20 lines described above was isolated from young leaves as described by Bernatzki and Tanksley (1986, Theor. Appl. Genet. **72**, 314-321). The typical yield was 50 - 100 µg DNA per gram of fresh leaf material. Template DNA for AFLP analysis with the enzyme combination *EcoRI*-*MseI* was

prepared as described by Zabeau and Vos (European Patent Application, EP 0534858), and is described briefly below:

0.5 µg of tomato DNA was incubated for 1 hour at 37 °C with 5 units *Eco*RI and 5 units *Mse*I in 40µl 10 mM Tris.HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA. Next 10 µl of a solution containing 5 pMol *Eco*RI-adapters, 50 pMol *Mse*I-adapters, 1 unit T4 DNA-ligase, 1 mM ATP in 10 mM Tris.HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA was added, and the incubation was continued for 3 hours at 37°C. The adapters are depicted below:

The structure of the *Eco*RI-adapter was:

```
5'-CTCGTAGACTGCGTACC
      CATCTGACGCATGGTTAA-5'
```

The structure of the *Mse*I-adapter was:

```
5'-GACGATGAGTCCTGAG
      TACTCAGGACTCAT-5'
```

Adapters were prepared by adding equimolar amounts of both strands; adapters were not phosphorylated. After ligation, the reaction mixture was diluted to 500 µl with 10 mM Tris.HCl, 0.1 mM EDTA pH 8.0, and stored at -20°C. The diluted reaction mixture is further referred to as template DNA.

AFLP reactions

The primers used for the AFLP screening are depicted below:

*Eco*RI-primers:	5'-GACTGCGTACCAATTCNNN-3'
*Mse*I-primers:	5'-GATGAGTCCTGAGTAANNN-3'

The N's in the primers indicate that this part of the primers was variable. In the AFLP screening all 64 possible primers were used for both the *Eco*RI- and *Mse*I-primer. This gave a total of 64 x 64 combinations of *Eco*RI- and *Mse*I-primers, is 4096 primer combinations. All 4096 primer combinations were used in the AFLP screening for *I-2* linked AFLP markers. The AFLP reactions were performed in the following way:

AFLP reactions employed a radio-actively labelled *Eco*RI-primer and a non-labelled *Mse*I-primer. The *Eco*RI-primers were end-labelled using (γ-^{33}P)ATP and T4 polynucleotide kinase. The labelling reactions were performed in 50 µl 25 mM Tris.HCl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM spermidine.3HCl using 500 ng oligonucleotide primer, 100 µCi (γ-^{33}P)ATP and 10 units T4 polynucleotide kinase. For AFLP analysis 20 µl reaction mixture were prepared containing 5 ng labelled *Eco*RI-primer (0.5 µl from the labelling reaction mixture), 30 ng *Mse*I-primer, 5 µl template-DNA, 0.4 units Taq-polymerase, 10 mM Tris.HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.2 mM of all 4 dNTPs. AFLP reactions were performed using the following cycle profile: a 30 seconds DNA denaturation step at 94 °C, a 30 seconds annealing step (see below), and a 1 minute extension step at 72 °C. The annealing temperature in the first cycle was 65 °C, was subsequently reduced each cycle by 0.7 °C for the next 12 cycles, and was continued at 56 °C for the remaining 23 cycles. All amplification reactions were performed in a PE-9600 thermocycler (Perkin Elmer Corp., Norwalk, CT, USA).

Gel analysis of AFLP reaction products

After amplification, reaction products were mixed with an equal volume (20 µl) of formamide dye (98% formamide, 10 mM EDTA pH 8.0, and bromo phenol blue and xylene cyanol as tracking dyes). The resulting mixtures were heated for 3 minutes at 90°C, and then quickly cooled on ice. 2 µl of each sample was loaded on a 5% denaturing (sequencing) polyacrylamide gel (Maxam and Gilbert, 1980, Methods in Enzymology **65**, 499-560). The gel matrix was prepared using 5% acrylamide, 0.25% methylene bisacryl, 7.5 M urea in 50 mM Tris/50 mM Boric acid/1 mM

EDTA. To 100 ml of gel solution 500 µl of 10% APS and 100 µl TEMED was added and gels were cast using a SequiGen 38 x 50 cm gel apparatus (Biorad Laboratories Inc., Hercules, CA, USA). Sharktooth combs were used to give 97 lanes on the SequiGen gel units. 100 mM Tris/100 mM Boric acid/2 mM EDTA was used as running buffer. Electrophoresis was performed at constant power, 110 Watts, for approximately 2 hours. After electrophoresis, gels were fixed for 30 minutes in 10% acetic acid dried on the glass plates and exposed to Fuji phospho image screens for 16 hrs. Fingerprint patterns were visualized using a Fuji BAS-2000 phospho image analysis system (Fuji Photo Film Company Ltd, Japan).

AFLP screening for linked markers

The template DNAs of the 20 tomato lines were pooled in the following way:

resistant pool 1:	tomato lines 1 - 5
resistant pool 2:	tomato lines 6 - 10
susceptible pool 3:	tomato lines 11 - 15
susceptible pool 4:	tomato lines 16 - 20

An AFLP screening was performed using all possible 4096 *Eco*RI-*Mse*I primer combinations on the 4 pools. The aim was to identify AFLP markers present in both resistant pools, and absent in both sensitive pools. AFLP gels contained the AFLP fingerprints of 24 primer combinations of the 4 pools, giving a total of 171 gels. Additional gels were run to reanalyse unsuccessful AFLP reactions and to confirm candidate markers. A total of 18 AFLP markers were identified present in both resistant pools and absent in both susceptible pools: these markers were referred to as candidate *I-2* linked markers.

Next AFLP reactions were performed to determine the presence of all 18 candidate markers on the 20 individual tomato lines. All markers appeared to be present in the 10 resistant lines and absent in the 10 susceptible lines. These 18 markers were named EM01 to EM18 and were used in further studies to map the *I-2* Fusarium resistance gene. The primer combinations required to identify markers

EM01 to EM18 are depicted in Table 1. In the column with the primer combinations, "*Eco*RI-" refers to the sequence 5'-GACTGCGTACCAATTC-3' and "*Mse*I-" refers to the sequence 5'-GATGAGTCCTGAGTAA-3'. For example, marker EM06 can be identified using the *Eco*RI-primer having the following sequence: 5'-GACTGCGTACCAATTCAGA-3', and the *Mse*I-primer having the following sequence: 5'-GATGAGTCCTGAGTAATCT-3'.

TABLE 1

marker	primer combination containing selective extensions (NNN)
EM01	*Eco*RI-AAA / *Mse*I-AGG
EM02	*Eco*RI-CCA / *Mse*I-TCA
EM03	*Eco*RI-CTC / *Mse*I-GCT
EM04	*Eco*RI-GAG / *Mse*I-GTC
EM05	*Eco*RI-TCT / *Mse*I-AAG
EM06	*Eco*RI-AGA / *Mse*I-TCT
EM07	*Eco*RI-CTT / *Mse*I-AAG
EM08	*Eco*RI-CAA / *Mse*I-GCT
EM09	*Eco*RI-GTC / *Mse*I-GTC
EM10	*Eco*RI-TGT / *Mse*I-AAT
EM11	*Eco*RI-TAG / *Mse*I-AGC
EM12	*Eco*RI-TGC / *Mse*I-AAG
EM13	*Eco*RI-TAA / *Mse*I-ACC
EM14	*Eco*RI-TGC / *Mse*I-AGA
EM15	*Eco*RI-CTT / *Mse*I-ATG
EM16	*Eco*RI-CAT / *Mse*I-AGT
EM17	*Eco*RI-CTC / *Mse*I-AGC
EM18	*Eco*RI-CGG / *Mse*I-CAC

EXAMPLE 3: GENETIC MAPPING OF THE TOMATO *I-2* LOCUS

Development of genetic material

Two *F. oxysporum* f.sp. *lycopersici* race 2 resistant tomato (*Lycopersicon esculentum*) lines have both been crossed with the susceptible tomato lines GCR210 and GCR508. Both resistant lines, DR9 and RZ5, are breeding lines from two seed companies, De Ruiter Zonen and Rijk Zwaan, respectively. The susceptible tomato line GCR210 is homozygous for the recessive morphological

marker gene *a* (anthocyaninless). The susceptible line GCR508 is homozygous for the recessive morphological marker gene *sub* (subtilis) (Stevens and Rick, 1986, in: The Tomato Crop, Atherton & Rudich edit., Chapman and Hall, p. 35-109). Both susceptible lines were obtained from the Institute of Horticultural Research (Littlehampton, United Kingdom).

The dominant *I-2* gene (conferring resistance to *F. oxysporum* f.sp. *lycopersici* 2), and recessive genes *a* and *sub* have all been mapped to the long arm of chromosome 11 of tomato (Stevens and Rick, 1986, in: The Tomato Crop, Atherton & Rudich edit., Chapman and Hall, p. 35-109): *I-2* on position 85 of this chromosome, *a* on position 68 (17 centiMorgans (cM) from *I-2*), and *sub* on position 89 (4 cM from *I-2*).

The crosses made are depicted below:

Cross 1:	DR9 (+ +, *I-2 I-2*)	x GCR210 (*a a*, *i-2 i-2*)
Cross 2:	RZ5 (+ +, *I-2 I-2*)	x GCR210 (*a a*, *i-2 i-2*)
Cross 3:	DR9 (*I-2 I-2*, + +)	x GCR508 (*i-2 i-2*, *sub sub*)
Cross 4:	RZ5 (*I-2 I-2*, + +)	x GCR508 (*i-2 i-2*, *sub sub*)

F_1 plants from all crosses were selfed for generating F_2 seeds. The resulting F_2 populations will segregate for resistance to *F. oxysporum* f.sp. *lycopersici* 2 and for the morphological markers.

Selection of recombinants in the region containing *I-2*

3000 F_2 seeds from the crosses 1 and 2 with GCR210 (containing *a*) were germinated. After 10 to 14 days, the seedlings were divided in two groups: one with purple-coloured hypocotyls (wild type seedlings) and the other with green hypocotyls (seedlings homozygous *a*). 330 wild type seedlings and 565 anthocyaninless seedlings were inoculated with *F. oxysporum* f.sp. *lycopersici* race 2. After three to four weeks the plants were scored for resistance/susceptibility. The results are shown below:

phenotype	resistant	susceptible
wild type	278	52
anthocyaninless	169	396

The recombinants are susceptible wild type plants and resistant anthocyaninless plants.

1500 F_2 seeds from the crosses 3 and 4 with GCR508 (containing *sub*) were germinated. After 10 to 14 days, the seedlings were inoculated with *F. oxysporum* f.sp. *lycopersici* race 2. During the assay the plantlets were phenotypically scored for growth habit, normal or 'subtilis', and for resistance/susceptibility. The results are shown below:

phenotype	resistant	susceptible
wild type	913	38
subtilis	18	351

The recombinants are susceptible wild type plants and resistant 'subtilis' plants.

F_2 plants were grown in the greenhouse for seed set (by induced or spontaneous selfing). F_3 seeds were obtained from most of the resistant recombinant F_2 plants and from a few of the susceptible recombinants. These F_3 lines were tested for resistance/ susceptibility in order to check the phenotype of the F_2 plants. Twenty to 30 seedlings of each F_3 line were inoculated with *F. oxysporum* f.sp. *lycopersici* 2 and evaluated as described in Example 1. The F_3 progenies of the resistant

recombinants did segregate 3:1 for resistance, while those of the susceptible recombinants scored all susceptible.

Screening of AFLP markers in F_2 recombinants

An AFLP analysis with the 18 identified *I-2* markers EM01 to EM18 has been performed on 187 resistant recombinant plants, 10 susceptible recombinant plants, and 39 control plants (12 resistant and 27 susceptible) as described in Example 2. The following results were obtained:

- The markers EM01, EM02, EM04, EM05, EM06, EM07, EM08, EM09, EM10, EM11, EM13, EM14, EM15 and EM17 were present in all 199 resistant F_2 plants and absent in all 37 susceptible plants; these markers are closely linked to the *I-2* gene.
- The markers EM03, EM12 and EM16 were present in all resistant plants and absent in the susceptible plants with the exception of one anthocyaninless plant.
- The marker EM18 is present in all resistant plants except for seven anthocyaninless plants, and absent in all susceptible plants.

From these data it could be concluded that the *I-2* locus is flanked by marker EM18 at one end of the DNA segment comprising the *I-2* resistance gene and on the other end by the markers EM03, EM12 and EM16. All of the remaining markers completely co-segregate with the *I-2* resistance gene based on the analysis of recombinants of crosses as described above.

EXAMPLE 4: CONSTRUCTION AND SCREENING OF A TOMATO YAC-LIBRARY

Material

The tomato line *Lycopersicon esculentum* E22 (Enza Zaden) homozygous for the *I-2* locus, was used as source DNA to construct a YAC-library. Protoplasts were

isolated from the leaves of *in vitro* shoots which were two to three weeks old as described by Van Daelen *et al.* (1989, Plant Mol. Biol. **12**, 341-352).

Viable protoplasts (concentration of 50 million protoplasts per ml) were collected and mixed with an equal volume of agarose (1%, Seaplaque, FMC BioProducts, Rockland, Maine, USA) to form a plug. The protoplasts embedded into the plugs were lysed with lysis mix (0.5 M EDTA, 1% N-Laurylsarcosinate and 1 mg/ml proteinase K, pH= 8.0). After lysis, the plugs were stored at 4 °C in storage buffer (fresh lysis mix) until used. Approximately 3 million protoplasts per plug, to obtain about 4.5 µg of chromosomal DNA were used for further studies. Plasmid pYAC4 containing an unique *Eco*RI cloning site was used as cloning vector and the yeast strain AB1380 was used as a host (Burke *et al.*, 1987, Science **236**, 806-812).

YAC library construction

High molecular weight DNA isolation, partial digestion with *Eco*RI in the presence of *Eco*RI methylase, ligation of vector arms to genomic DNA, size selection by pulsed field gel electrophoresis and transformation of the yeast host was performed as described by Burke *et al.* (1987, Science **236**, 806-812) and Larin *et al.* (1991, Proc. Natl. Acad. Sci. USA **88**, 4123-4127).

All standard manipulations were carried out as described in Molecular cloning: a laboratory manual by Sambrook *et al.*, (1989, Cold Spring Harbor Laboratory Press).

3840 clones with a average insert size of 520 kb, which corresponds to 2.2 genome equivalents were finally obtained and the individual clones were stored in 40 96-wells microtiter plates containing 75 µl YPD solution (1% yeast extract, 2% peptone and 2% dextrose).

Screening YAC library

To reduce the number of samples handled, the cells of one 96-well microtiter plate was pooled (a plate pool) and used for DNA isolation as described by Ross *et al.* (1991, Nucleic Acids Res., **19**, 6053). The 2.2 genome equivalent tomato YAC library consists of 40 96-wells microtiter wells and as a result DNA of the 40 plate

pools were screened with the AFLP-markers EM01, EM12 and EM18 (see Example 2) using the AFLP-protocol as described in Example 2. One positive plate pool out of the 40 was identified with all three AFLP-markers. Subsequently, a secondary screening of the 96 individual YAC clones was employed to find the correct address of the YAC or YACs. One individual YAC clone was identified, designated 1/546, and subsequently analyzed with the remaining AFLP markers. As expected, all of identified markers EM01 to EM18 were present on this YAC-clone since the flanking markers and one co-segregating marker were used in the screening. The size of the YAC-clone was determined by Pulse-field gel electrophoretic (PFGE) analysis using contour-clamped homogeneous electric field (CHEF; Chu *et al.* 1986, Science, **235**, 1582-1585) and appeared to be 750 kb.

EXAMPLE 5: CONSTRUCTION OF A PHYSICAL MAP OF YAC 1/546 AND LOCATION OF THE AFLP MARKERS

YAC 1/546 was subjected to partial digestion with increasing concentration of the restriction enzymes *Sgr*A1, *Rsr*II, *Sfi*I and *Bss*HII. The samples were fractionated by PFGE, transferred to a Gene Screen Plus membrane (DuPont NEN, Boston, MA, USA) and assayed by hybridization using end-adjacent sequence probes according to the protocol for indirect end-label mapping as described by Burke *et al.* (1987, Science **236**, 806-812).

In Figure 1 a schematic representation is given of the physical map of YAC 1/546. To determine the position of the various AFLP markers on the physical map, the AFLP markers were used as hybridization probes on the partial digests of YAC 1/546. Therefore, the AFLP-marker fragment was excised from the dried gel and eluted by means of diffusion in a buffer containing 0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA (pH=8.0), 0.1% SDS, re-amplified with the PCR primers used in the AFLP reaction and labelled with ^{32}P according to the random primer method of Feinberg and Vogelstein (1983, Anal. Biochem. **132**, 6-10).

It appeared that all of the AFLP-markers were located on a *Sgr*A1 fragment, at a distance of 255 kb of the left arm of the YAC until the first *Sgr*A1 site.

EXAMPLE 6: CONSTRUCTION OF A COSMID LIBRARY OF YAC 1/546

Material

The binary cosmid vector pJJ04541 is a derivative of pJJ1881 (Jones *et al.*, 1992, Transgenic Research **1**, 285-297) and is based on plasmid pRK290 containing the tetracyclin resistance gene for selection in *Escherichia coli* and *Agrobacterium tumefaciens*. Into the unique *Eco*RI site of pRK290, T-DNA carrying sequences (LB; left border repeat, RB signifies the right border repeat) that flank

- the cos site of bacteriophage lambda
- the neomycin phosphotransferase gene (Beck *et al.*, 1982, Gene **19**, 327-336) whose expression is driven by the cauliflower mosaic virus 35S promoter sequence (Odell *et al.*, 1984, Mol. Gen. Genet. **223**, 369-378), and
- the pBluescript (Stratagene, La Jolla, CA, USA) polylinker sequence.

The size of pJJ04541 amounts 29 kb and is shown schematically in Figure 2. The cloning capacity of this binary cosmid vector, using phage lambda packaging extracts is within the range of 9 to 24 kb.

Library construction

Total DNA of the *Saccharomyces cerevisae strain* AB1380 containing YAC 1/546 was isolated using zymolyase to make protoplasts according to Green and Olsen (1990, Proc. Natl. Acad. Sci. USA **87**, 1213-1217).

An aliquot was analyzed on PFGE and appeared to have a size of ≥100 kb.

Approximately 15 µg of this DNA was partially digested with *Sau*3A generating molecules with an average size of 15-25 kb. Subsequently, this sample was centrifugated through a 10-35% sucrose gradient for 22 hours, 22.000 rpm at 20 °C in a Beckman SW41 rotor. 0.5 ml fractions were collected using a needle pierced through the bottom of the centrifuge tube. An aliquot of these fractions was

analyzed on a 0.7% agarose gel. The fractions containing DNA molecules with a size of ≈20 kb were pooled and concentrated by ethanol precipitation.

Subsequently, the cohesive ends were partially filled-in with dATP and dGTP using the strategy of partial filling of 5'-extensions of DNA produced by type II restriction endonuclease as described by Korch (1987, Nucleic Acids Res. **15**, 3199-3220) and Loftus *et al.* (1992, Biotechniques **12**, 172-176).

The binary cosmid vector pJJ04541 was digested completely with *Xho*I and the linear fragment was partially filled-in with dTTP and dCTP as described by Korch (1987, Nucleic Acids Res. **15**, 3199-3220).

The 20-kb fragments were ligated to the cosmid vector and transduced to *E. coli* strain XL1-Blue MR (Stratagene, La Jolla, CA, USA) using phage lambda Gigapack II XL packaging extracts (Stratagene, La Jolla, CA, USA) as recommended by the manufacturers. Selection was performed on LB (1% bacto-tryptone, 0.5% bacto-yeast extract and 1% NaCl, pH 7.5) agar plates containing 10 mg/l of tetracyclin. A bank of approximately 250.000 cosmid clones was made from 2-3 µg of size fractionated yeast DNA.

Subsequently, these transformants were stored into the wells of microtiter plates (96-wells, 100 µl of LB medium containing 10 mg/l of tetracyclin). Replicas of the 96-well grid of cosmid clones in microtiter plates were stamped onto Gene Screen Plus membrane filters (DuPont NEN, Boston, MA, USA) and allowed to grow into colonies on media. Colony hybridization using ^{32}P-labelled *Sgr*A1 fragment revealed positive cosmids. Of about 10.000 colonies approximately 150 positive cosmid clones were identified.

EXAMPLE 7: DETAILED PHYSICAL MAP OF THE 255 SgrA1 FRAGMENT AND LOCATION OF THE AFLP MARKERS

Construction of a cosmid contig of the 255 kb *Sgr*A1 fragment

Standard techniques for growth and manipulation of cosmids in *E. coli* were followed (Sambrook *et al.* in Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press). Cosmid DNA was isolated by alkaline lysis using

the method as described by Ish-Horowicz *et al.* (1981, Nucl. Acids Res. **9**, 2989-2997). Approximately 500 ng was used for template preparation and the primers in the amplification of restriction fragments were the *Eco*RI-primer 5'-GACTGCGTACCAATTC-3' having no selective nucleotide and the *Mse*I-primer 5'-GATGAGTCCTGAGTAA-3' having no selective nucleotide as described in Example 2. The *Eco*RI-primer was labelled at the 5' end and each of the 150 DNAs was amplified using the described primer set. The DNA fingerprints contained about 8 to 20 amplified fragments. Sets of DNA samples containing amplified fragments of identical size were selected and were rerun on polyacrylamide gels as described in Example 2 until a contiguous array of all the amplified fragments throughout the *Sgr*A1 fragment was obtained. The final fingerprint of the cosmid contig is shown in Figure 3.

Restriction fragment amplification is an excellent tool for contig building of cosmids, however the level of overlap between two adjacent cosmids cannot be determined using the enzyme combination as described above. Therefore, the DNA samples of the cosmid contig were digested with *Eco*RI and *Hin*dIII followed by Southern blot analysis according to Southern, J. Mol. Biol. **98**, 503-515. ^{32}P-labelled *Sgr*AI fragment was used as a probe and the sizes of the overlapping fragments between two adjacent cosmid DNAs was determined making use of size markers. Adjacent cosmids having an overlap of at least 5 kb were used for complementation analysis.

Fine mapping of the 18 AFLP markers on the 255 kb *Sgr*AI fragment

The indirect end-label mapping technique as described by Burke *et al.* (1987, Science **236**, 806-812), for the restriction enzymes MluI and SalI and using the left-end probe (see described in example 5) was performed to construct a more precise physical map of the 255 kB SgrAI fragment containing the *I-2* resistance gene. Pulse field gel electrophoresis was performed under conditions that maximise resolution in the range up to 100 kb.

To position the AFLP markers EM01 to EM18 on the physical map of the 255 kb *Sgr*A1 fragment two types of hybridization analysis were performed. First, the

AFLP markers were used as probes on the *Mlu*I and *Sal*I partial digests and secondly the AFLP markers were used as probes on DNAs of the cosmid contig using standard hybridization techniques as described by Sambrook *et al.* (in Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press).

A physical fine map for the enzymes *Mlu*I and *Sal*I of the *Sgr*A1 fragment and the position of the 18 AFLP markers EM01 to EM18 is shown in Figure 4.

EXAMPLE 8: TRANSFORMATION

Transfer of cosmids to *Agrobacterium tumefaciens*

The cosmid clones were introduced in *Agrobacterium tumefaciens* through conjugative transfer in a tri-parental mating with helper strain HB101 (pRK2013) essentially according to Deblaere *et al.* (1987, Methods in Enzymology **153**, 277-292). *E.coli* were grown in LB medium (1% bacto-tryptone, 0.5% bacto-yeast extract and 1% NaCl, pH 7.5) supplemented with 5 mg/l tetracyclin at 37°C. The helper strain HB101 (pRK2013) was grown under identical conditions in LB medium supplemented with 100 mg/l kanamycin sulphate. *Agrobacterium tumefaciens* strain C58C1Rif^R (pGV3101) (Van Larebeke *et al.*, 1974, Nature, **252**, 169-170) was grown in LB medium supplemented with 100 mg/l rifampicin at 28°C. Overnight cultures were collected by centrifugation at 4000 r.p.m. for 5 min and resuspended in LB medium without any supplements, and 0.5 ml each of the *Agrobacterium* culture, the helper strain culture and a cosmid strain culture were mixed and plated on LB agar plates without antibiotics. After overnight incubation at 28°C, the mixture was plated on LB agar plates containing 100 mg/l rifampicin and 5 mg/l tetracyclin to select for single transconjugant *Agrobacterium* colonies in serial passages through selective agar plates.

Characterization of A. tumefaciens transconjugants

Small-scale cultures were grown from selected colonies and grown in LB medium containing 10 mg/l tetracyclin. Plasmid DNA was isolated by alkaline lysis using

the method as described by Ish-Horowicz et al. (1981, Nucl. Acids Res. **9**, 2989-2997). and digested with *Bgl*II using standard techniques. In addition, restriction fragment amplification on miniprep DNA of *A. tumefaciens* was performed using the enzyme combination *Eco*RI/*Mse*I and primers having no selective nucleotide as described in Example 7. Subsequently, the *Bgl*II restriction enzyme pattern as well as the DNA fingerprint of the *A. tumefaciens* transconjugant were compared with those of miniprep DNA of the *E. coli* strain containing the cosmid. Only those *A. tumefaciens* transconjugants harbouring a cosmid with the same DNA pattern as the corresponding *E. coli* culture were used to transform a susceptible tomato line.

Transformation of a susceptible tomato line

Seeds of the susceptible tomato line 52201 (Rijk Zwaan) were surface-sterilized in 2% sodium hypochlorite for 10 min, rinsed three times in sterile distilled water, and placed on germination medium (consisting of half-strength MS medium according to Murashige and Skoog (1962, Physiol. Plant. **15,** 473-497), with 1% (w/v) sucrose and 0.8% agar) in glass jars or polypropylene culture vessels. They were left to germinate for 8 days. Culture conditions were 25°C, a photon flux density of 30 $\mu mol.m^{-2}.s^{-1}$ and a photoperiod of 16 /24 h.

Transformation of tomato was performed according to Koornneef *et al.* (1987, In: Tomato Biotechnology, 169-178, Alan R. Liss, Inc.), and is described briefly below. Eight day old cotyledon explants were precultured for 24 h in Petri dishes containing a feeder layer of *Petunia hybrida* suspension cells plated on MS20 medium (culture medium according to Murashige and Skoog (1962, Physiol. Plant. **15,** 473-497) with 2% (w/v) sucrose and 0.8% agar) supplemented with 10.7 µM α-naphthaleneacetic acid and 4.4 µM 6-benzylaminopurine. The explants were then infected with the diluted overnight culture of *Agrobacterium tumefaciens* containing the cosmid clone AA12, DD2, CC14, A52, B22, A55, CC16, A44, A29, CC3, AA2, AA9, BB8 or the cosmid vector pJJ04541 for 5-10 min, blotted dry on sterile filter paper and cocultured for 48 h on the original feeder layer plates. Culture

conditions were as described above. Overnight cultures of *Agrobacterium tumefaciens* were diluted in liquid MS20 medium (medium according to Murashige and Skoog (1962, Physiol. Plant. **15,** 473-497), with 2% (w/v/) sucrose, pH 5.7) to an $O.D._{600}$ of 0.8.

Following the cocultivation, the cotyledon explants were transferred to Petri dishes with selective medium consisting of MS20 supplemented with 4.56 µM zeatin, 67.3 µM vancomycin, 418.9 µM cefotaxime and 171.6 µM kanamycin sulphate, and cultured under the culture conditions described above. The explants were subcultured every 3 weeks onto fresh medium. Emerging shoots were dissected from the underlying callus and transferred to glass jars with selective medium without zeatin to form roots. The formation of roots in a medium containing kanamycin sulphate was regarded as an indication of the transgenic nature of the shoot in question. Truly transgenic regenerants were propagated *in vitro* by subculturing the apical meristem and auxiliary buds into glass jars with fresh selective medium without zeatin.

Seed production

Of every individual transformant, one or two clones were kept *in vitro*, while the remaining shoots were potted in soil, hardened off in a relative humidity >85% at 20 °C and a photoperiod of 16/24h for 7-10 days, and transferred to the greenhouse for seed production. Mature tomato plants were grown at 25-32 °C in a ventilated greenhouse with supplementary illumination from Son-T lamps (Philips, Eindhoven, The Netherlands). They were watered daily, treated against mildew and Trialeurodes infestation and vibrated to ensure good seed set when flowering. Mature fruits were harvested, the seeds extracted, cleaned in 1% HCl for 1 h, rinsed with water and dried.

EXAMPLE 9: COMPLEMENTATION ANALYSIS

Identification of cosmids with the resistance gene by screening for resistance in transformed plants

The transformed plants (R_0 plants) of the transformation experiments were grown in the greenhouse for seed set as described in Example 8. For each of the cosmids AA12, DD2, CC14, A52, B22, A55, CC16, A44, A29, CC3, AA2, AA9, BB8 or pJJ04541 twenty R_0 plants were grown. R_1 lines of at least nine R_0 plants of each cosmid were tested for disease symptoms, except for pJJ04541 in which case four plants were tested, in order to identify cosmids with the *I-2* resistance gene. Twenty to 25 seedlings of each R_1 line were inoculated and evaluated as described in Example 1. In total 144 R_1 lines of the above mentioned 14 different cosmid transformations have been tested; 13 cosmids contain tomato insert DNA and one cosmid, pJJ04541, is without insert DNA. The results are shown in Table 2. 128 transgenic R_0 plants appeared susceptible, because their R_1 lines were completely or nearly completely (at least 80% of the seedlings) obviously diseased. Sixteen R_0 plants are resistant, because their R_1 lines were segregating. Most of the R_1 lines segregated in a ratio of about 3:1, resistant and susceptible seedlings. The remaining R_1 lines had a small number of susceptible seedlings (segregation ratios of 5:1 and 10:1), indicating the presence of more than one copy of the insert in the corresponding R_0 plants. The 16 resistant R_0 plants had been derived from transformations with one of the following cosmids: A52, B22 and A55; these three cosmids overlap with each other. These identified cosmids were used for further molecular analysis.

TABLE 2

cosmid	R_0 plants	
	resistant	susceptible
AA12	0	10
DD2	0	11
CC14	0	10
A52	4	7
B22	8	3
A55	4	9
CC16	0	15
A44	0	10
A29	0	11
CC3	0	10
AA2	0	9
AA9	0	10
BB8	0	9
pJJ04541	0	4

Molecular analysis of the transformed plants with a resistant phenotype

To demonstrate that the resistant phenotype of transgenic plants, which were transformed with one of the overlapping cosmids A52, B22 and A55, is determined by the genomic insert present in the various cosmids, an AFLP analysis with the AFLP markers EM04, EM06 and EM14 was performed. The AFLP markers EM14 and EM06 were positioned on plasmids A52 and B22 whereas EM06 and EM04 were positioned on cosmid A55 (see Figure 5)

DNA was isolated from young leaves of resistant as well as susceptible R_1 plants derived from a resistant R_0 plant, as described by Bernatzki and Tanksley (1986, Theor. Appl. Genet. **72**, 314-321). Selective restriction fragment amplification was

performed with the primer combinations identifying the markers EM14 and EM06 for the R_0 plants transformed with cosmids A52 and B22 and with the primer combinations identifying the markers EM06 and EM04 for the R_0 plants transformed with cosmids A55. The DNA fingerprints obtained showed in both cases and for both markers the presence of the markers in the resistant plants and the absence of the markers in the susceptible plants indicating that the three identified overlapping cosmids A52, B22 and A55 contain the *I-2* resistance gene.

Confirmation of the resistant phenotype in second generations of transformed plants

Additional genetic evidence for the presence of the *I-2* resistance gene on cosmids A52, B22 and A55 has been obtained in the next generation of the R_1 plants. Resistant plants of the R_1 lines, transformed with cosmid A52, B22 or A55 that segregated in a ratio of 3:1 (or higher) were grown in the greenhouse for seed set as described in Example 8. All plants were selfed and most of the plants were backcrossed with the susceptible line 52201, the originally transformed tomato genotype in this invention, as the male parent. Of each resistant R_0 plant at least two R_2 lines (resulting from selfing of R_1 plants) and, if available, two R_1BC lines (resulted from R_1 backcrossed with 52201) were tested for Fusarium 2 resistance, in order to confirm inheritance of the introgressed *I-2* resistance gene. Seedlings of the R_2 lines and R_1BC lines were inoculated and evaluated as described in Example 1.

At least half of the seedlings of each tested line appeared to be resistant. Most R_2 lines were segregating 3:1 or were completely resistant, while most R_1BC lines were segregating 1:1 or were completely resistant. The results indicate that R_1 plants were either heterozygous for the *I-2* resistance gene (segregations of 3:1 in the R_2 and 1:1 in the R_1BC) or homozygous for the *I-2* resistance gene (all seedlings in the R_2 and R_1BC were resistant).

EXAMPLE 10: PHYSICAL MAP OF THE OVERLAPPING COSMID CLONES CONTAINING THE *I-2* RESISTANCE GENE

Since the size of the cosmid inserts is within the range of 13-24 kb, standard techniques such as single, double and partial digestion analysis with restriction enzymes which are present in the polylinker sequence of the cosmid vector were performed as described by Sambrook *et al.* (in: Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press). A physical map of cosmid A52, B22 and partially A55, was constructed and the overlap between the three cosmids giving rise to a resistant phenotype (A52,B22 and A55) with respect to the adjacent cosmid clone CC16 revealing a susceptible phenotype, was determined and is depicted in Figure 5.

The insert size in cosmids A52 and B22 could be calculated and amount 23 and 17 kb, respectively. It appeared that the plant insert of cosmid B22 completely fits within cosmid A52. The minimal DNA segment containing the *I-2* resistance gene was defined by the left-end of cosmid A55 until the right-end of cosmid B22 encompassing a region of approximately 8 kb in size.

EXAMPLE 11: NUCLEOTIDE SEQUENCE AND DEDUCED AMINO ACID SEQUENCE OF THE *I-2* RESISTANCE GENE

Various fragments of the 8 kb DNA segment of cosmid B22 were subcloned into the *E. coli* vector pBluescript (Stratagene, La Jolla, CA, USA) using standard techniques as described by Sambrook *et al.* (in: Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press) and used for sequence analysis making use of the Pharmacia Autoread Sequencing Kit and the Pharmacia LKB A.L.F. DNA Sequencer device (Pharmacia LKB, Uppsala, Sweden). The nucleotide sequence of 6.5 kb ranging from approximately 600 nucleotides upstream of a *Hin*dIII site until the 3' end of cosmid B22 was determined and is shown in Figure 6. A large open reading frame of 3798 nucleotides encoding a protein of 1266 amino acids could be deduced.

EXAMPLE 12 : TRANSCRIPT MAPPING

Transcript mapping studies were performed to map the 5' and 3' end of the *I-2* resistance gene and to determine whether the *I-2* resistance gene comprises any introns. The polymerase chain reaction to amplify parts of the transcripts from the *I-2* resistance gene was used for this purpose.

Total RNA from leaf tissue of the resistant tomato cultivar E22 was isolated using the hot phenol method as described by Sambrook *et al.* (in: Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press). Poly A+ RNA was isolated using biotinylated oligo(dT) bound to Dynabeads M-280 Streptavidin (DYNAL A.S., Oslo, Norway) according to the instructions of the manufacturer. A cDNA library was constructed using the Superscript RNase H Reverse Transcriptase cDNA kit from Life technologies, Inc. (Gaithersburg, MD, USA) and the protocol supplied by the manufacturer.

The cDNA was used as template for oligonucleotide primers in various PCR reactions. These primers were designed based on the nucleotide sequence provided in Figure 6, and six primer sets were used covering the coding sequence of the *I-2* resistance gene. In addition, 5' and 3' RACE products were obtained using the Marathon cDNA amplification kit from Clontech (Paolo Alto, CA, USA). Subsequently, the various PCR fragments obtained with the 6 internal primer sets, and the 5' and 3'-RACE fragments were cloned into the TA cloning vector pCRII (Invitrogen Corporation, San Diego, CA, USA) and sequenced using standard protocols. The nucleotide sequences obtained were aligned with the 6.5 kb genomic sequence and three intron sequences could be deduced. An intron of 86 nucleotides was located just upstream the ATG initiation codon from nucleotide position 1703 to 1788 (Figure 6). The 5' end of the largest transcript detected with the Marathon cDNA amplification kit, after analysis of 16 individual transformants, maps at nucleotide position 1597. Hence, we conclude that the transcriptional initiation site of the *I-2* resistance gene is positioned at or upstream of nucleotide 1597.

Two other intron sequences with a size of 399 and 82 nucleotides were located in the 3' untranslated region at nucleotide position 5628 to 6026 and 6093 to 6174,

respectively. The 3' end of the transcript was mapped at nucleotide position 6491 based on the sequence the 3' end RACE product. Hence, we conclude that the transcriptional termination site of the *I-2* resistance gene is positioned at or immediatly downstream of nucleotide 6491. A putative poly-adenylation signal (AAUAAA) is located at nucleotide position 6406-6411, approximately 80 nucleotides upstream of the transcriptional termination site and the stretch of U residues. The various introns and the position of the transcriptional initiation and termination site are depicted in Figure 6.

The DNA sequence of the internal PCR products derived from cDNA as template indicated that the *I-2* resistance gene does not contain any introns within the coding region.

EXAMPLE 13: DNA SEGMENT COMPRISING THE *I-2* RESISTANCE GENE.

Based on the physical map of the three overlapping cosmid clones comprising the I2-resistance gene, it could be deduced that the DNA segment comprising the *I-2* resistance gene was defined by the left end of cosmid A55 until the right end of cosmid B22 encompassing a region of approximately 8 kb in size. One large open reading frame with a size of 3798 nucleotides encoding a protein of 1266 amino acids could be deduced from this 8 kb region. Part of this 8 kb DNA segment was retransformed into the susceptible tomato line 52201.

As starting vector for subcloning of the DNA segment, the cointegrate vector pKG1505 was used. This vector is a derivative of plasmid pGV1500, the prototype cointegrate vector described by Deblaere *et al.* (1987, Methods in Enzymology, **153**, 277-292). The vector is based on the common cloning vector pBR322. It contains a streptomycin/spectinomycin resistance gene to be used as selection marker in *Agrobacterium tumefaciens* and maintenance of the cointegrate structure. The vector contains the left (LB) and right (RB) border repeat sequences of the octopine T_L-DNA. Vector pKG1505 differs from pGV1500 in that it contains no residual T-DNA between the border repeat sequences, but a *nos-nptII-nos*

cassette for kanamycin resistance in plants instead and a synthetic polylinker sequence.

Using standard techniques for the construction of recombinant DNA molecules as described by Sambrook *et al.* (in: Molecular cloning: a laboratory manual, 1989, Cold Spring Harbor Laboratory Press) plasmid pKG6016 (Figure 7) was constructed. Plasmid pKG6016 comprises a 6.2 kb *ScaI/SalI* fragment from cosmid B22 comprising the coding sequence of the *I-2* resistance gene and a 1.3 kb upstream and 1.1 downstream DNA region.

For this construction, a 3.8 kb *BamHI/SalI* fragment from cosmid B22 was cloned into pKG1505 resulting in plasmid pKG6014. This 3.8 kb DNA segment contains the second half of the *I-2* resistance gene and the 1.1 kb 3' untranslated region.

Moreover, a 2.4 kb *ScaI/BamHI* fragment from cosmid B22 was cloned into the *EcoRV/BamHI* sites of pBluescript (Stratagene, La Jolla, CA, USA) resulting in plasmid pKG6015. Subsequently, pKG6015 was cut with *XhoI* and *BamHI* and the 2.4 kb DNA segment containing the first half of the *I-2* resistance gene including a 1.3 kb upstream promoter region was introduced into pKG6014, resulting into plasmid pKG6016.

Plasmid pKG6016 was introduced into *Agrobacterium tumefaciens* strain C58C1Rif^R (pGV2260) (Deblaere *et al.*, 1987, Methods in Enzymology, **153**, 277-292) through conjugative transfer in a tri-parental mating with helper strain HB101 (pRK2013) as described before. Selection was performed on plates containing 300 µg/ml of spectinomycin and 100 µg/ml of streptomycin.

The *Agrobacterium tumefaciens* transconjugants were characterized by Southern blot analysis of chromosomal DNA using the Sm gene as a probe. Therefore, chromosomal DNA was isolated according to the method described by Lichtenstein and Draper (in: DNA cloning, 1985, volume II, 67-119, D.M. Glover edit., IRL Press), digested with the restriction enzymes *EcoRI* or *HindIII* and blotted onto a nitrocellulose membrane (Gene Screen Plus, DuPont NEN, Boston, MA, USA). Those transconjugants containing one copy of the integrative plasmid were

selected for transformation to the susceptible tomato line 52201 using the protocol as described before.

Individual transformants were transferred to the greenhouse of which fifteen R_0 plants were grown for seed production. R_1 lines of six R_0 plants were tested for disease symptoms. Twenty to 30 seedlings of each R_1 line were inoculated with *Fusarium oxysporum* f.sp. *lycopersici* race 2 and evaluated: wilting plants were considered to be susceptible, whereas not wilting plants were considered to be resistant. The observations indicated that the DNA segment is involved in the resistance.

CLAIMS

1. A nucleic acid whose DNA sequence is at least part of the DNA sequence provided in figure 6 or any DNA sequence homologous thereto.

2. The nucleic acid of claim 1 wherein said DNA sequence homologous to the DNA sequence of figure 6 is capable, when transferred to a host plant also liable of being rendered resistant by the DNA sequence of figure 6, of also rendering it resistant to Fusarium 2.

3. The nucleic acid of claim 1 which is capable, when transferred to a host plant, which is susceptible to a plant pathogen, of rendering said host plant resistant to said plant pathogen.

4. A nucleic acid according to claim 1 wherein said DNA sequence corresponds to a coding sequence starting at nucleotide 1798 and ending at nucleotide 5598 or any DNA sequence homologous thereto.

5. A nucleic acid according to claim 1 wherein said DNA sequence corresponds to a promoter sequence located 5' upstream of nucleotide 1798 or any DNA sequence homologous thereto.

6. A nucleic acid according to claim 1 wherein said DNA sequence corresponds to a sequence starting at nucleotide 464 and ending at nucleotide 6658 or any DNA sequence homologous thereto.

7. A nucleic acid of claim 1 wherein said DNA sequence corresponds to at least part of the genomic insert present in cosmid B22, and preferably corresponds to the overlapping genomic DNA sequence between cosmid B22 and cosmid A55, or any DNA sequence homologous thereto.

8. A nucleic acid of claim 7 wherein said overlapping genomic DNA sequence is defined by the left end of the genomic insert present in cosmid A55 and the right end of the genomic insert present in cosmid B22.

9. A recombinant DNA construct comprising a nucleic acid according to any of claims 1-8.

10. A recombinant DNA construct of claim 9 in which said nucleic acid is under control of a promoter which is functional in a plant cell, said promoter being

either endogenous or exogenous to said plant cell, and effective to control the transcription of said DNA sequence in such plant cells.

11. A recombinant DNA construct of claim 10 in which said promoter corresponds to a promoter sequence located 5' upstream of nucleotide 1797 as provided in figure 6, or any DNA sequence homologous thereto.

12. A vector suitable for transforming plant cells comprising a DNA construct according to any of claims 9-11.

13. Plasmid pKGI2-B22 as deposited under number CBS 546.95.

14. Plasmid pKGI2-A55 as deposited under number CBS 820.96.

15. Bacterial cells comprising a vector or plasmid according to any of claims 12-14.

16. Recombinant plant genome comprising, incorporated thereinto, a DNA construct according to any of claims 9-11.

17. Plant cells comprising a DNA construct according to any of claims 9-11.

18. Plant comprising plant cells according to claim 17.

19. Plant according to claim 18 which has a reduced susceptibility to Fusarium 2.

20. Plant according to claim 19 wherein said plant is tomato and wherein said Fusarium 2 is *Fusarium oxysporum* f.sp. *lycopersici* race 2.

21. Seed comprising a DNA construct according to any of claims 9-11.

22. The recombinant plant genome of claim 16, in a plant cellular environment.

23. Process for obtaining plants having reduced susceptibility to a fungus, comprising the following steps:

i) inserting into the genome of a plant cell a DNA construct according to any of claims 9-11,

ii) obtaining transformed plant cells,

iii) regenerating from said transformed plant cells genetically transformed plants, and

iv) optionally, propagating said plants.

24. Process according to claim 23 further comprising selecting transformed plants having reduced susceptibility to said fungus.

25. Process according to claim 23 or 24 wherein said fungus is a soil born fungus, and preferably a wilt inducing fungus.

26. Process according to any of claims 23-25 wherein said plant is tomato and wherein said fungus is *Fusarium oxysporum* f.sp. *lycopersici* race 2.

27. Process for protecting plants in cultivation against fungal infection, which comprises:

i) providing the genome of plants with a DNA construct according to any of claims 9-11, and

ii) growing said plants.

28. Process for isolating a nucleic acid according to claim 1-8, comprising the following steps:

i) screening a genomic or cDNA library of a plant with a DNA sequence according to claim 1-8,

ii) identifying positive clones which hybridize to said DNA sequence,

iii) isolating said positive clones.

29. The process of claim 28 wherein said library originates from a first plant and the DNA sequence belongs to a second plant.

30. Process of selective restriction fragment amplification for identifying a nucleic acid according to claim 1-8 using primer combinations identifying at least one of the AFLP markers EM01 to EM18.

31. The process of claim 30 wherein said primer combination identifies AFLP marker EM06.

32. An oligonucleotide comprising a DNA sequence which corresponds to at least part of the nucleic acid according to claim 1-8.

33. The oligonucleotide of claim 32, which is of a size sufficient to hybridize selectively to the DNA sequence of any of claims 1 to 8 under stringent hybridization conditions.

34. An oligonucleotide according to claim 32 wherein said DNA sequence corresponds to the sequence starting at nucleotide 3470 and ending at nucleotide 3565.

35. An oligonucleotide according to claim 32 wherein said DNA sequence is located at the 3'end, and preferably corresponds to the sequence 5'-AATTCAGA-3', which can prime the synthesis of DNA.

36. An oligonucleotide according to claim 32 wherein said DNA sequence is located at the 3'end, and preferably corresponds to the sequence 5'-TAATCT-3' which can prime the synthesis of DNA.

37. A primer combination comprising a first oligonucleotide according to claim 35 and a second oligonucleotide according to claim 36.

38. Diagnostic kit comprising at least one oligonucleotide according to any of claims 32-36.

39. Diagnostic kit comprising a primer combination according to claim 37.

40. Process for detecting the presence or absence of a DNA sequence according to claim 1-8, particularly in a plant DNA using a diagnostic kit according to claim 38 or 39.

41. A polypeptide having an amino acid sequence having the sequence provided in figure 6 or coded by the corresponding homologous sequence according to claim 1 or 2.

42. Process for the identification of elicitor molecules using the polypeptide according to claim 41 as a receptor molecule.

43. A RNA having a ribonucleic acid sequence of a transcript of part or all of the DNA sequence of claim 1 or 2.

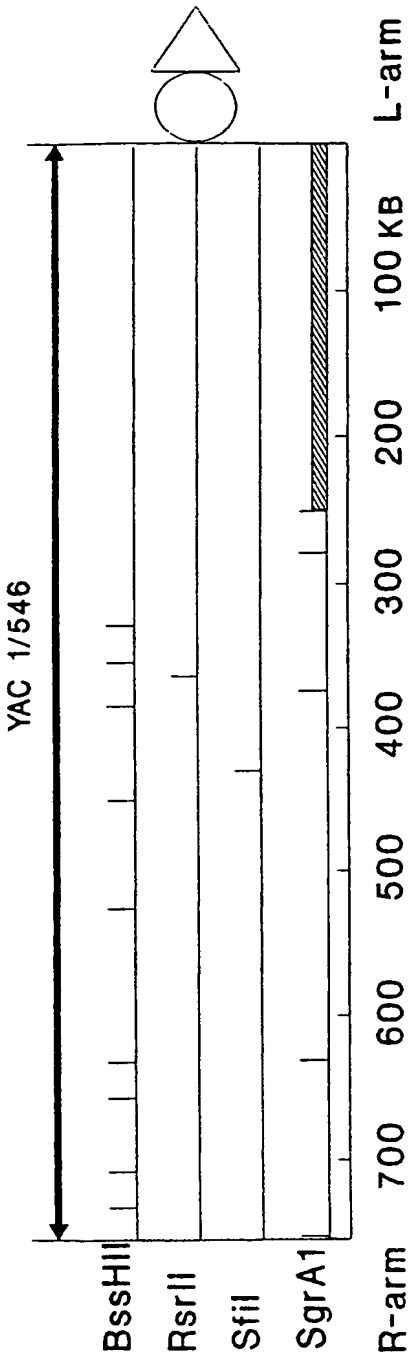
YAC 1/546
BssHII
RsrII
SfiI
SgrA1
R-arm
700
600
500
400
300
200
100 KB
L-arm

FIGURE 1

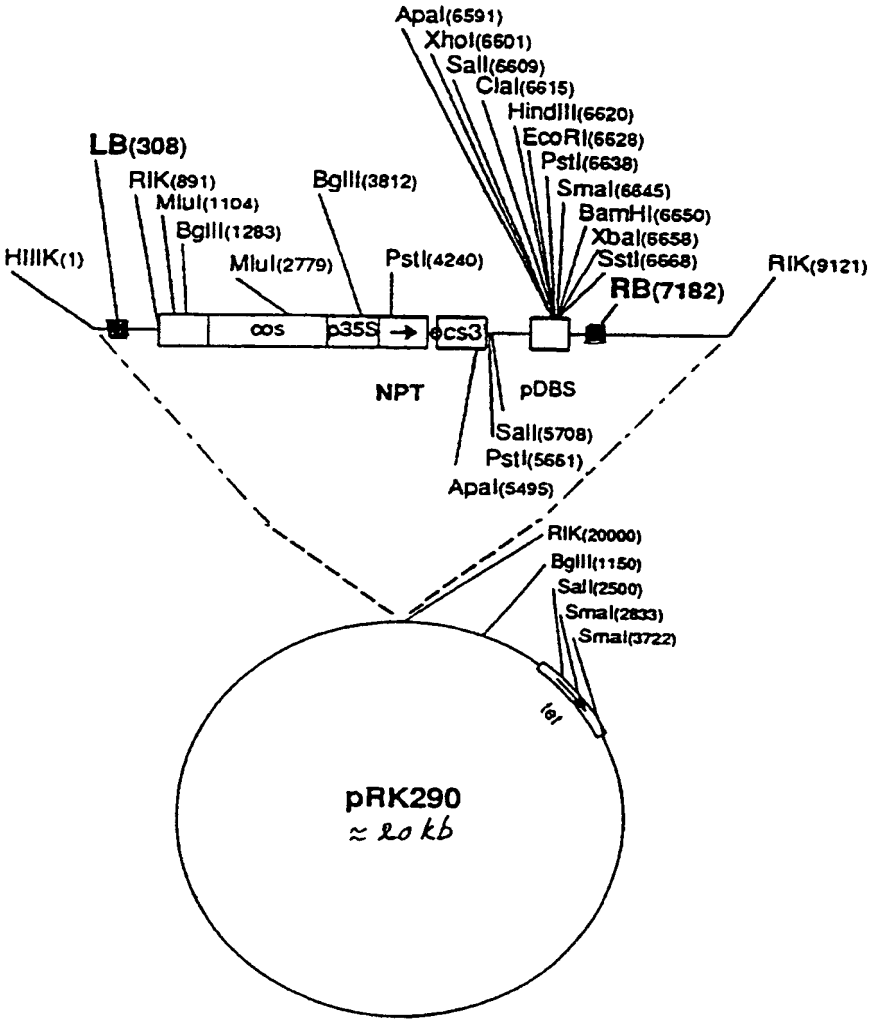
ApaI(6591)
XhoI(6601)
SalI(6609)
ClaI(6615)
HindIII(6620)
EcoRI(6628)
PstI(6638)
SmaI(6645)
BamHI(6650)
XbaI(6658)
SstI(6668)
RB(7182)
LB(308)
RIK(891)
MluI(1104)
BglII(1283)
MluI(2779)
BglII(3812)
PstI(4240)
HIIIK(1)
RIK(9121)
cos
p35S
ocs3
NPT
pDBS
SalI(5708)
PstI(5661)
ApaI(5495)
RIK(20000)
BglII(1150)
SalI(2500)
SmaI(2833)
SmaI(3722)
ori
pRK290
≈ 20 kb

FIGURE 2

QA31001L N [..]

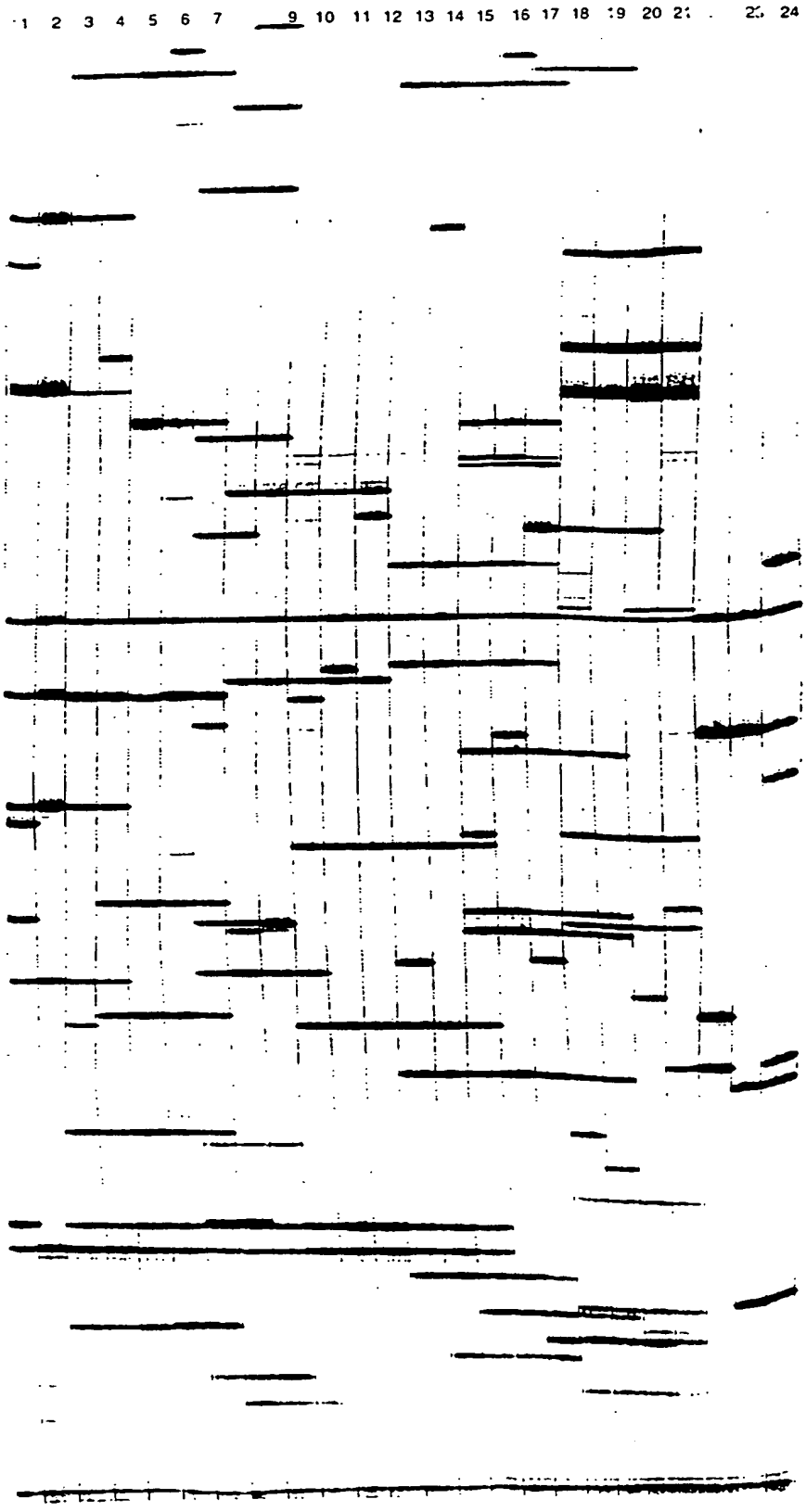

FIGURE 3

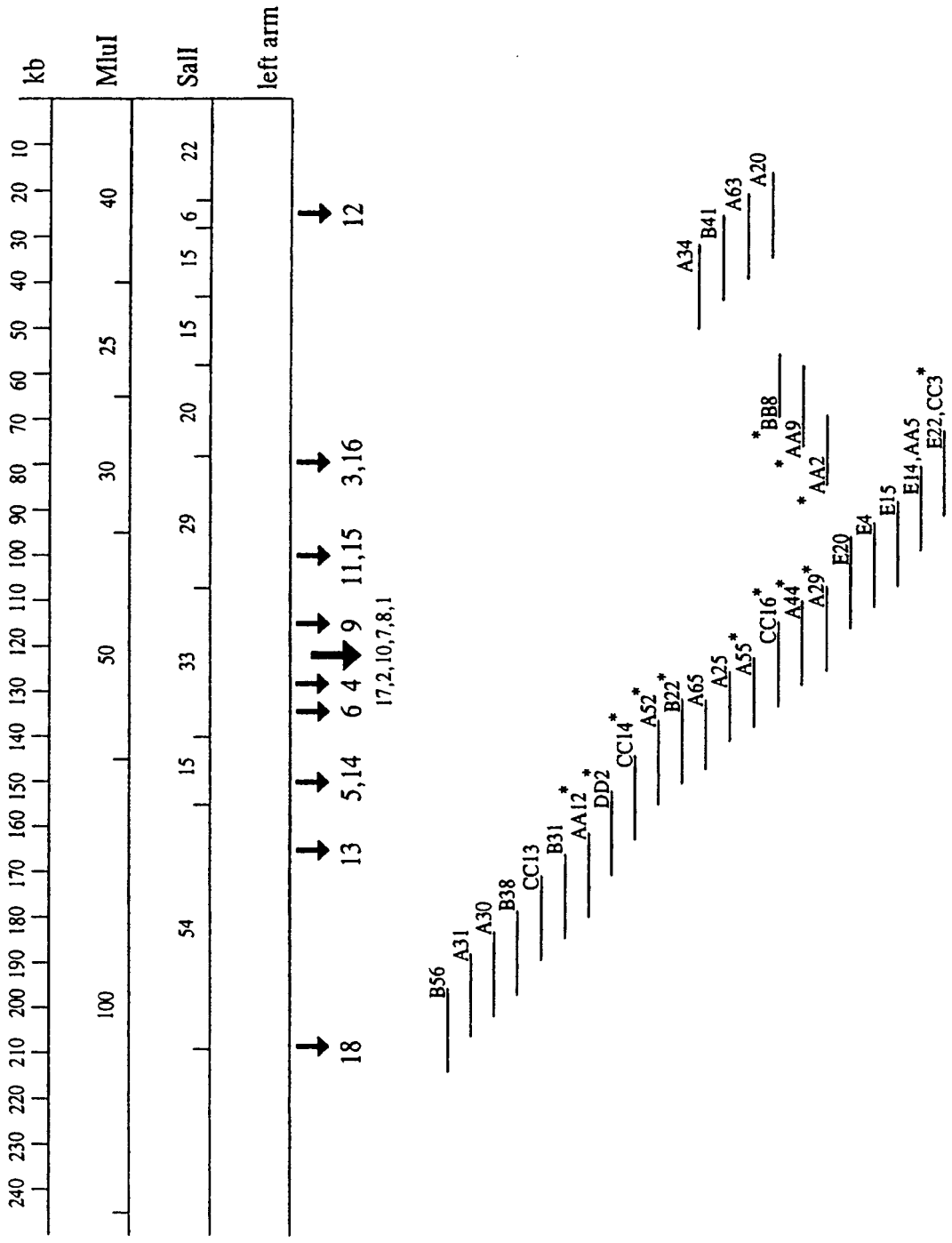
kb
MluI
SalI
left arm
10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240
100
50
30
40
54
15
33
29
20
15
15
6
22
18
13
5,14
6 4
9
11,15
3,16
12
17,2,10,7,8,1
B56
A31
A30
B38
CC13
B31
AA12*
DD2*
CC14*
A52*
B22*
A65
A25
A55*
CC16*
A44*
A29*
E20
E4
E15
E14,AA5
E22,CC3*
AA2*
AA9*
BB8*
A34
B41
A63
A20

FIGURE 4

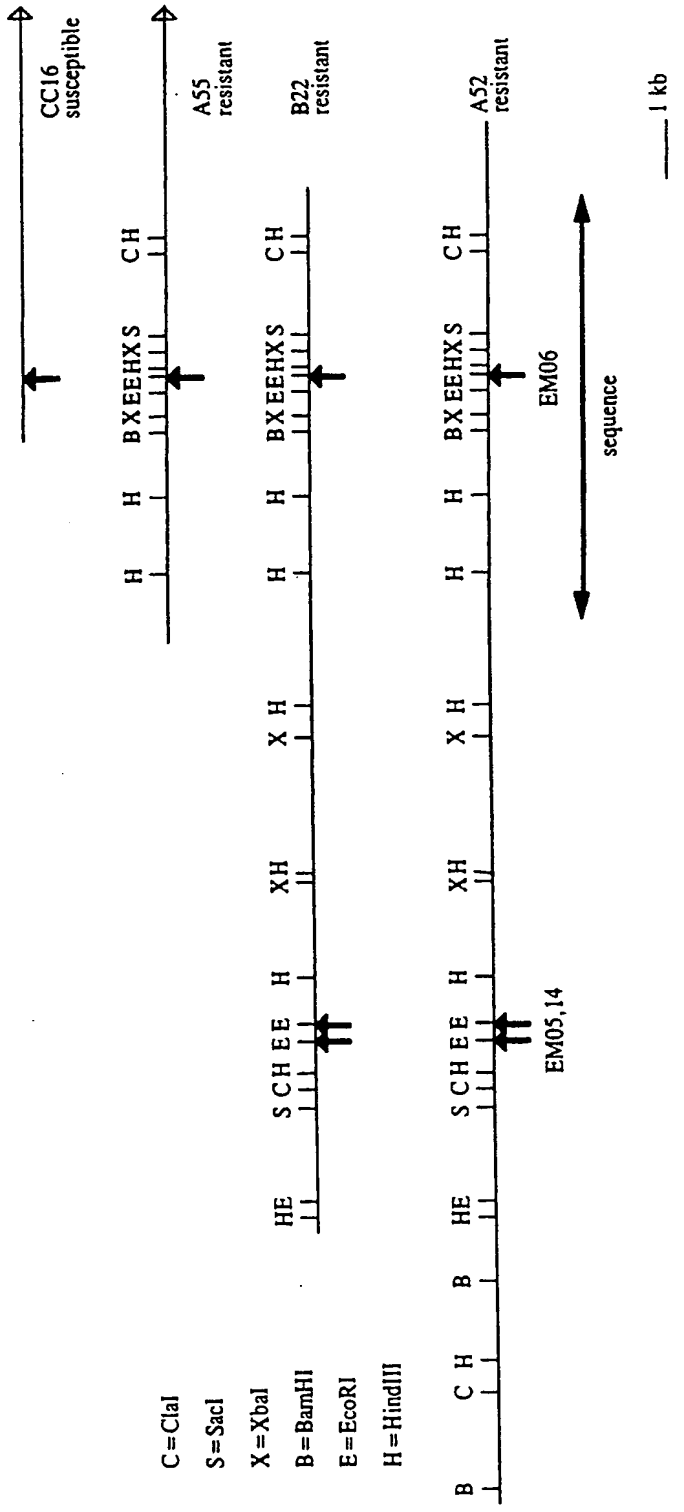
CC16 susceptible
A55 resistant
B22 resistant
A52 resistant
EM06
sequence
EM05,14
1 kb
C=ClaI
S=SacI
X=XbaI
B=BamHI
E=EcoRI
H=HindIII

FIGURE 5

```
5'-AAAAAAGCAGCTTTAAAAAAAGTACTTTKGAAAGGKGCTGAAACTTATTTTTTGAAATAA    60
   GCAGTTATGTGTTTGGAAWAAAAGTGCTGAAGTTGCTATGTCAAACATGAAAAGGGRAAA   120
   AATGGAAGAAAGAGWTGTTAGGGTTATGTCGTAATTTGGAGATTGTATAAAAATATTAAG   180
   GGCAAAAAAAAATAAAAATGTGTCAACTTAAAACAGCTTATAAGCTAAAAGTTAAAAGCTG   240
   GGGTAGAGGTGTTTTTTTTTTTTTTAGCTTATAAGTTGTTTTAAGTTGACCACATTTTTA   300
   TTTTTKTTGCCCTTAATATTTTTATACAATCTCAAAATTACGACATAACCCTAACATCTT   360
   TTTCTCCCATTTTTTCCTTTTCACGTTTGACATAGCAACTTCAGCACTTTTATCCAAACA   420
   CATAACTGCTTATTTTAAAAATAAGTTTCAGCACTTTCAAAAGTACTTTTTTAAAGCTGC   480
   TTTTATTAAGCCCATCCAAACGGGCCCTAAAATTGCTAATGTTTGCTCTTTCTATTCTCA   540
   AACTCCGTAATATTTAAGAAAATTTGCTAATGATAGGTCACTTTTAACACTAAATAATTA   600
   TAAATTGGGTAGAAATTTATTTATCATTTTAAGCTTTTTTAATTTTGAGTCTTCTCCCTA   660
   ATTAAGACCCTTCCCCTCTTGCTTCAATTATTTAACTGAATAGTCTTTGTCTTATTGTTG   720
   GGTGAAAGTCTGTCTTCTTGTTAGGTACTAAGTCCTACAATAATATCAATAATTTGCTAT   780
   GGAGAAAAAAAATATTATAGGAGAAAAATAATTAATTTTAATTCATGAATATGTCTTAATA   840
   TGCAACTCATTTTGCTTATATATATCAAATTAAACTCTGTTCCTTTAACTTTTTCCTATG   900
   AAGATACATTTTAATTTATTTGATGAGGTTAGTTTTGAAATTTATATTATAATAATGAAA   960
   TGATATAACTTAAAAGAAGTTGTTTGATATCTTATCAGAATCATGCAGGTACTCATAATA  1020
   TAAGAAATAATTATGATGAAATTTATATATGTTTTATGCAGAGATTTATTACGCATTGTT  1080
   TACTTGGGTTATGTATTACTTATTTCATCTTTTATCAGAATGTAAAATTATCATTCAATA  1140
   AGAAATCCAATTCTGTTAAATTCAAAATACAAACAATAACATTTTCAAGACCGATTTTTT  1200
   GCCCAAGAATATACAGTAAACATATTTATGATATGGTAGGTCTCTTTAGTAATTGACCAA  1260
   CAAGGATTGTGGTGGAGTGGGAAATACTCTTTAATACTTCACCAAGAGGTCTCCAATTTG  1320
   AGCCCCTGAATACGAAATCGTCTTTGTTAGTATATACCCTAACCTAATACAAAAATTAGT  1380
   ATATTAGCCTTNACAGCTAAAATCTTTGTGACCTGTAAGTCACGCGAGGACAAATTTACC  1440
   GTAACACCAACTTATTCATGATATAATTGTCCCTTTTAGCACGGTAATAATGAGGTGGGT  1500
   AGAAATTTATTACTTGAGGGCCCTTTCTACACCCACCCTTATTCTCTTGCTTCAATTATT  1560
   GAATTGAAGAAGTAATGAAAAAACAGACTCCATTGGATAAAGGACAGTTTGCAAACACAG  1620
   CTGTAACAATTTAGAGCACTAGCAAAATAGAGAGAGTTTTGAGAGAAATTTTTGTTTGCA  1680
   AATTACTCTTAACCTTCAGCAG GTAAAATAAAGTTCTTAACTGAGACTATTTGAAGATAT  1740
   ATTTTGTTAAAGAATCATTTTGTGTGTTTCCTTGTTTTGCTTTTGCAGATTTGAGAAATG  1800
                                                          M   1
   GAGATTGGCTTAGCAGTTGGTGGTGCATTTCTCTCCTCAGCTTTGAATGTTCTGTTTGAT  1860
    E  I  G  L  A  V  G  G  A  F  L  S  S  A  L  N  V  L  F  D   21
   AGGCTTGCTCCTAACGGTGATCTGCTCAACATGTTTCGGAAGCATAAGGATCATGTTAAG  1920
    R  L  A  P  N  G  D  L  L  N  M  F  R  K  H  K  D  H  V  K   41
   CTCTTAAAGAAGCTGAAAATGACTTTGCGTGGTATTCAGATTGTGCTAAGTGATGCAGAG  1980
    L  L  K  K  L  K  M  T  L  R  G  I  Q  I  V  L  S  D  A  E   61
   AATAAGCAAGCATCAAATCCATCTGTGAGAGACTGGCTTAATGAGCTTCGAGATGCTGTC  2040
    N  K  Q  A  S  N  P  S  V  R  D  W  L  N  E  L  R  D  A  V   81
   GACTCTGCTGAAAATTTAATAGAAGAAGTCAATTATGAAGCTTTGAGGCTTAAGGTGGAA  2100
    D  S  A  E  N  L  I  E  E  V  N  Y  E  A  L  R  L  K  V  E  101
   GGTCAGCATCAGAATTTTTCAGAAACAAGCAACCAGCAAGTAAGTGATGATTTTTTCCTT  2160
    G  Q  H  Q  N  F  S  E  T  S  N  Q  Q  V  S  D  D  F  F  L  121
   AACATAAAGGACAAGCTGGAAGACACTATTGAAACATTAAAGGATTTGCAAGAGCAAATT  2220
    N  I  K  D  K  L  E  D  T  I  E  T  L  K  D  L  Q  E  Q  I  141
   GGTCTCCTTGGCTTAAAGGAGTATTTTGATTCCACGAAACTAGAAACTAGAAGACCTTCA  2280
    G  L  L  G  L  K  E  Y  F  D  S  T  K  L  E  T  R  R  P  S  161
   ACTTCTGTGGATGATGAATCTGATATCTTTGGTAGGCAGAGCGAAATAGAGGATTTGATT  2340
    T  S  V  D  D  E  S  D  I  F  G  R  Q  S  E  I  E  D  L  I  181
   GACCGTCTATTGTCTGAAGGTGCAAGTGGGAAAAAGCTGACAGTAGTTCCTATCGTTGGA  2400
    D  R  L  L  S  E  G  A  S  G  K  K  L  T  V  V  P  I  V  G  201
   ATGGGCGGCCAGGGCAAGACAACACTTGCTAAAGCCGTATACAATGATGAGAGGGTGAAG  2460
    M  G  G  Q  G  K  T  T  L  A  K  A  V  Y  N  D  E  R  V  K  221
   AATCATTTTGATTTGAAAGCGTGGTATTGCGTTTCTGAAGGATTTGATGCTTTGAGAATA  2520
    N  H  F  D  L  K  A  W  Y  C  V  S  E  G  F  D  A  L  R  I  241
   ACAAAAGAATTACTCCAAGAAATTGGCAAATTTGACTCGAAGGATGTCCACAACAATCTT  2580
    T  K  E  L  L  Q  E  I  G  K  F  D  S  K  D  V  H  N  N  L  261
```

FIGURE 6a

```
AACCAGCTTCAAGTCAAATTGAAGGAAAGTTTGAAGGGAAAGAAGTTCCTTATTGTTTTG  2640
 N  Q  L  Q  V  K  L  K  E  S  L  K  G  K  K  F  L  I  V  L  281
GATGATGTGTGGAATGAAAATTACAACGAGTGGAATGACTTGAGAAATATTTTTGCACAA  2700
 D  D  V  W  N  E  N  Y  N  E  W  N  D  L  R  N  I  F  A  Q  301
GGAGATATAGGAAGTAAGATCATTGTGACGACACGCAAAGACAGTGTTGCCTTGATGATG  2760
 G  D  I  G  S  K  I  I  V  T  T  R  K  D  S  V  A  L  M  M  321
GGAAATGAGCAAATTCGCATGGGCAATTTGTCTACCGAAGCCTCTTGGTCTTTATTTCAA  2820
 G  N  E  Q  I  R  M  G  N  L  S  T  E  A  S  W  S  L  F  Q  341
AGACATGCATTTGAAAACATGGATCCTATGGGACATCCGGAACTTGAAGAGGTCGGAAGA  2880
 R  H  A  F  E  N  M  D  P  M  G  H  P  E  L  E  E  V  G  R  361
CAAATTGCAGCCAAGTGCAAAGGACTGCCCTTAGCTCTGAAGACGCTCGCTGGCATGTTA  2940
 Q  I  A  A  K  C  K  G  L  P  L  A  L  K  T  L  A  G  M  L  381
CGCTCCAAATCAGAGGTTGAAGAGTGGAAACGTATTTTGAGAAGTGAAATATGGGAGCTG  3000
 R  S  K  S  E  V  E  E  W  K  R  I  L  R  S  E  I  W  E  L  401
CCACACAATGACATATTACCAGCGTTGATGTTGAGCTACAATGATCTTCCCGCACATTTA  3060
 P  H  N  D  I  L  P  A  L  M  L  S  Y  N  D  L  P  A  H  L  421
AAGCGATGCTTTTCTTTTTGTGCAATATTTCCTAAAGATTATCCATTTAGGAAAGAACAA  3120
 K  R  C  F  S  F  C  A  I  F  P  K  D  Y  P  F  R  K  E  Q  441
GTTATTCATCTATGGATTGCCAATGGTCTCGTACCAGTGAAAGATGAAATAAATCAAGAT  3180
 V  I  H  L  W  I  A  N  G  L  V  P  V  K  D  E  I  N  Q  D  461
TTAGGCAACCAATACTTTCTAGAGTTGAGATCAAGATCATTATTTGAAAAGGTCCCAAAT  3240
 L  G  N  Q  Y  F  L  E  L  R  S  R  S  L  F  E  K  V  P  N  481
CCTTCTAAAAGGAACATAGAGGAATTATTCCTTATGCATGACCTTGTCAATGATTTAGCC  3300
 P  S  K  R  N  I  E  E  L  F  L  M  H  D  L  V  N  D  L  A  501
CAACTTGCATCTTCAAAACTTTGTATCAGGTTAGAAGAGAGCCAAGGATCTCATATGTTG  3360
 Q  L  A  S  S  K  L  C  I  R  L  E  E  S  Q  G  S  H  M  L  521
GAACAATGTCGGCACTTATCTTATTCAATAGGATTTAATGGTGAGTTTAAGAAATTGACA  3420
 E  Q  C  R  H  L  S  Y  S  I  G  F  N  G  E  F  K  K  L  T  541
CCCCTCTACAAATTGGAGCAGTTGAGGACATTGCTTCCGATACGTATTGAATTCAGATTG  3480
 P  L  Y  K  L  E  Q  L  R  T  L  L  P  I  R  I  E  F  R  L  561
CACAATCTAAGCAAGAGGGTGTTGCATAACATACTGCCTACACTAAGATCCTTGAGGGCC  3540
 H  N  L  S  K  R  V  L  H  N  I  L  P  T  L  R  S  L  R  A  581
CTATCATTCTCTCAATACAAGATTAAGGAGTTGCCAAATGACTTGTTTACCAAATTAAAG  3600
 L  S  F  S  Q  Y  K  I  K  E  L  P  N  D  L  F  T  K  L  K  601
CTCCTCAGATTTTTGGATATTTCTCGGACATGGATTACAAAGTTGCCGGATTCCATTTGT  3660
 L  L  R  F  L  D  I  S  R  T  W  I  T  K  L  P  D  S  I  C  621
GGATTATATAACTTGGAGACACTTCTCCTGTCATCTTGTGCTGATCTTGAGGAGCTACCG  3720
 G  L  Y  N  L  E  T  L  L  L  S  S  C  A  D  L  E  E  L  P  641
CTGCAGATGGAGAAGTTGATTAACTTGCGTCATCTTGACGTAAGCAACACTCGGCGCTTG  3780
 L  Q  M  E  K  L  I  N  L  R  H  L  D  V  S  N  T  R  R  L  661
AAGATGCCACTACATCTGAGCAGGTTGAAAAGCCTCCAAGTGTTGGTGGGACCCAAGTTT  3840
 K  M  P  L  H  L  S  R  L  K  S  L  Q  V  L  V  G  P  K  F  681
TTTGTAGATGGTTGGAGAATGGAAGATTTGGGTGAAGCACAAAACTTACATGGATCTCTA  3900
 F  V  D  G  W  R  M  E  D  L  G  E  A  Q  N  L  H  G  S  L  701
TCAGTTGTGAAGTTGGAAAATGTGGTTGATAGAAGGGAAGCTGTGAAGGCAAAGATGAGG  3960
 S  V  V  K  L  E  N  V  V  D  R  R  E  A  V  K  A  K  M  R  721
GAGAAGAATCATGTTGAGCAATTATCATTGGAGTGGAGTGAAAGTAGTATTGCCGACAAT  4020
 E  K  N  H  V  E  Q  L  S  L  E  W  S  E  S  S  I  A  D  N  741
TCACAAACAGAAAGTGACATACTTGATGAGCTATGCCCACATAAAAACATCAAAAAAGTC  4080
 S  Q  T  E  S  D  I  L  D  E  L  C  P  H  K  N  I  K  K  V  761
GAAATCAGTGGATATAGAGGGACAAACTTTCCCAATTGGGTAGCTGATCCTTTGTTTCTT  4140
 E  I  S  G  Y  R  G  T  N  F  P  N  W  V  A  D  P  L  F  L  781
AAGCTGGTGAATTTGTCTCTAAGAAACTGCAAGGACTGTTACTCCTTGCCAGCACTAGGA  4200
 K  L  V  N  L  S  L  R  N  C  K  D  C  Y  S  L  P  A  L  G  801
CAACTCCCTTGTTTGAAATTCCTTTCCGTTAAAGGGATGCATGGAATAAGAGTGGTGACG  4260
 Q  L  P  C  L  K  F  L  S  V  K  G  M  H  G  I  R  V  V  T  821
```

FIGURE 6b

```
GAAGAATTCTATGGCAGATTGTCCTCCAAAAAGCCTTTTAACTCTCTAGAGAAGCTTGAA  4320
 E  E  F  Y  G  R  L  S  S  K  K  P  F  N  S  L  E  K  L  E   841
TTTGAAGATATGACGGAGTGGAAGCAATGGCACGCACTAGGAATTGGAGAGTTCCCTACA  4380
 F  E  D  M  T  E  W  K  Q  W  H  A  L  G  I  G  E  F  P  T   861
CTTGAGAACCTTTCAATTAAAAATTGCCCTGAGCTCAGTTTGGAGATACCCATCCAATTT  4440
 L  E  N  L  S  I  K  N  C  P  E  L  S  L  E  I  P  I  Q  F   881
TCAAGTTTAAAAAGGTTAGAAGTTAGTGATTGTCCAGTTGTTTTTGATGATGCCCAACTG  4500
 S  S  L  K  R  L  E  V  S  D  C  P  V  V  F  D  D  A  Q  L   901
TTTAGATCCCAACTTGAGGCAATGAAGCAGATTGAGGAAATAGATATATGTGATTGTAAC  4560
 F  R  S  Q  L  E  A  M  K  Q  I  E  E  I  D  I  C  D  C  N   921
TCTGTTACCTCCTTTCCTTTTAGCATACTGCCAACTACCTTGAAGAGAATACAGATATCT  4620
 S  V  T  S  F  P  F  S  I  L  P  T  T  L  K  R  I  Q  I  S   941
CGTTGCCCAAAATTGAAATTGGAGGCGCCAGTTGGTGAGATGTTTGTGGAGTATTTGAGA  4680
 R  C  P  K  L  K  L  E  A  P  V  G  E  M  F  V  E  Y  L  R   961
GTGAATGATTGTGGTTGTGTAGATGATATATCACCTGAGTTTCTCCCAACAGCACGTCAA  4740
 V  N  D  C  G  C  V  D  D  I  S  P  E  F  L  P  T  A  R  Q   981
TTGAGTATTGAAAATTGCCAGAACGTTACTAGGTTTTTGATTCCTACTGCCACTGAAACT  4800
 L  S  I  E  N  C  Q  N  V  T  R  F  L  I  P  T  A  T  E  T  1001
CTCCGTATTTCGAATTGTGAGAATGTTGAAAAACTATCGGTGGCATGTGGAGGAGCGGCC  4860
 L  R  I  S  N  C  E  N  V  E  K  L  S  V  A  C  G  G  A  A  1021
CAGATGACGTCACTGAATATTTGGGGATGTAAGAAGCTCAAGTGTCTTCCAGAACTCCTT  4920
 Q  M  T  S  L  N  I  W  G  C  K  K  L  K  C  L  P  E  L  L  1041
CCATCTCTCAAGGAACTGCGTCTGTCTGATTGTCCAGAAATAGAAGGAGAATTGCCCTTC  4980
 P  S  L  K  E  L  R  L  S  D  C  P  E  I  E  G  E  L  P  F  1061
AATTTAGAAATACTCCGTATCATATATTGCAAGAAACTGGTGAATGGCCGAAAGGAGTGG  5040
 N  L  E  I  L  R  I  I  Y  C  K  K  L  V  N  G  R  K  E  W  1081
CATTTACAGAGACTCACAGAGTTATGGATCGATCATGATGGGAGTGACGAAGATATTGAA  5100
 H  L  Q  R  L  T  E  L  W  I  D  H  D  G  S  D  E  D  I  E  1101
CATTGGGAGTTGCCTTGTTCTATTCAGAGACTTACCATAAAGAATCTTAAAACATTAAGC  5160
 H  W  E  L  P  C  S  I  Q  R  L  T  I  K  N  L  K  T  L  S  1121
AGCCAACATCTCAAAAGCCTCACCTCTCTTCAATATCTATGTATTGAGGGTTATTTATCT  5220
 S  Q  H  L  K  S  L  T  S  L  Q  Y  L  C  I  E  G  Y  L  S  1141
CAGATTCAGTCACAAGGCCAGCTTTCCTCCTTTTCTCACCTCACTTCGCTTCAAACTCTA  5280
 Q  I  Q  S  Q  G  Q  L  S  S  F  S  H  L  T  S  L  Q  T  L  1161
CAAATCTGGAATTTCCTTAATCTCCAATCACTTGCTGAATCAGCACTGCCCTCCTCCCTC  5340
 Q  I  W  N  F  L  N  L  Q  S  L  A  E  S  A  L  P  S  S  L  1181
TCTCACCTGGAGATAGATGATTGCCCTAATCTCCAATCACTCTTCGAATCAGCACTGCCC  5400
 S  H  L  E  I  D  D  C  P  N  L  Q  S  L  F  E  S  A  L  P  1201
TCCTCCCTCTCTCAGCTGTTCATCCAGGATTGCCCTAATCTCCAATCCCTTCCATTTAAA  5460
 S  S  L  S  Q  L  F  I  Q  D  C  P  N  L  Q  S  L  P  F  K  1221
GGGATGCCCTCTTCCCTCTCTAAACTATCTATTTTCAATTGCCCATTGCTCACACCACTA  5520
 G  M  P  S  S  L  S  K  L  S  I  F  N  C  P  L  L  T  P  L  1241
CTAGAATTTGACAAGGGGGAATACTGGCCACAAATTGCTCATATTCCCATCATAAATATC  5580
 L  E  F  D  K  G  E  Y  W  P  Q  I  A  H  I  P  I  I  N  I  1261
GATTGGAAATATATTTAACAATTAAAACAAATGGCTCTCCAACTGATGTAAGCTATTCGT  5640
 D  W  K  Y  I  ---                                            1266
TACCCTTAGAAGCTTTTTATGATTCTATGTTTCTCATTGCTTATTGGTTTATGCTCTTAC  5700
CGTGTTTTAATTCACGTCTCAATTGCCACCATGTTTAATCGAAAGTTTTTAGTTCTTGTA  5760
ATCATCAACCATCCTATGTCACTAGAAATTTTGATAGGTAAAAGAGGTAGACAAAAAAGC  5820
TAAACATCTTTTTTCTTTCGTATAGCGACCAAACAACTACATTTTGATAGGTAAGGGCTA  5880
TAGATATACATTTGCAGGGTGTTAAACCAAGGAGTAAGAAAATCACTGTCTTCAGATATC  5940
TTCTCTTGCATATACTTTTGCAATTTTAAGCTACATTTTGAACTCATGTGTTGTTGCTAA  6000
CTTAAACATGTTTTGTGCTTAATCAGATGTGGATTTTGAAGAGCGAGTACGACAAGTCTG  6060
GTACATTAATTGTCCGTAGAAGTGTTTCTAAGGTGCTGCTGCTATTTTTACATCTGTTCC  6120
CGAGTTTTGTTTTTTTTTTAAATCTTTCCACTAAAGCTATTATGTCGTCCACAGTGAATT  6180
TTCAGGTCTGTTGTTATAGGCAAGTCTTTGAGATGCGACTATCAAAGAAGGGCGATTACA  6240
```

FIGURE 6c

```
ATCAGTGTACCGCTGAAACTATTTCATGTTTCCAGTGCAAGCCTCTTTTGTAAGTTGACA  6300
AACTCGATTAGTTAATATGTTTGGGACTCAACTAGTGGTTAGAGTACTCATTTTGTAAGA  6360
CTTGTGTACAGAAAATCAAATTAGAATTATAACTCGTGATGGTTGAATAAACTCTAAGAA  6420
GTACTGATATATTTTTTAGTGGATATGTTGTTTGCTCATTCGGTGTTTGATATCCACATT  6480
GGAGTCCAACTAAATTCGAATTTGCACAATCGAAGGAGCGGTGCTCCTGGCATGATTTTT  6540
TTCCCATTCTACGACTAGTGCTCCTAAATTCTAATTAAGCATAGAAAAATCTCAACTATC  6600
TCACCCAACTCATATCAGGATAGAGTATTCCCTGAGGAGGATTCCTTCAGTTACAAAA-3' 6658
```

FIGURE 6d

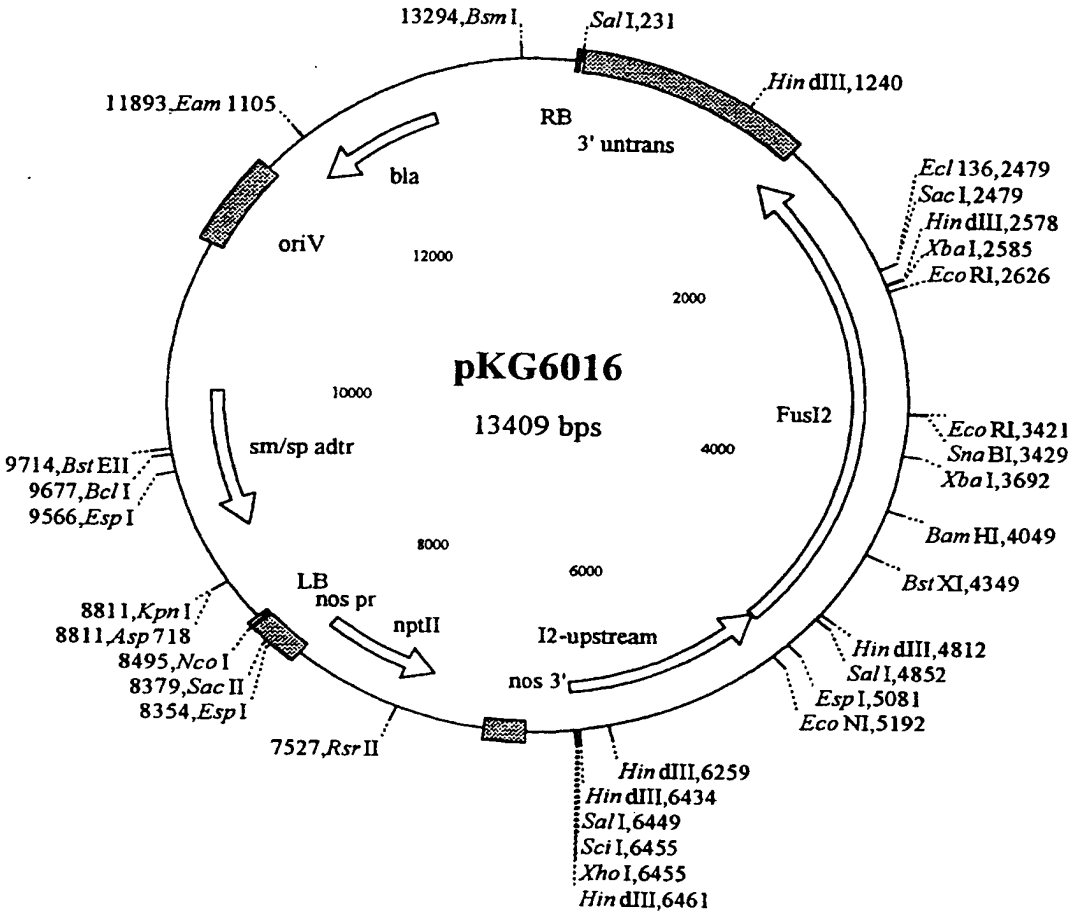
pKG6016
13409 bps
13294,BsmI
SalI,231
HindIII,1240
EclI36,2479
SacI,2479
HindIII,2578
XbaI,2585
EcoRI,2626
EcoRI,3421
SnaBI,3429
XbaI,3692
BamHI,4049
BstXI,4349
HindIII,4812
SalI,4852
EspI,5081
EcoNI,5192
HindIII,6259
HindIII,6434
SalI,6449
SciI,6455
XhoI,6455
HindIII,6461
7527,RsrII
8354,EspI
8379,SacII
8495,NcoI
8811,Asp718
8811,KpnI
9566,EspI
9677,BclI
9714,BstEII
11893,Eam1105
RB
3' untrans
bla
oriV
sm/sp adtr
LB
nos pr
nptII
I2-upstream
nos 3'
FusI2
2000
4000
6000
8000
10000
12000

FIGURE 7